Hutson
_76072_
_1652_

```
                GenCore version 4.5
        Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      August 29, 2002, 07:54:53 ; Search time 69.02 Seconds
                                          (without alignments)
                                          793.551 Million cell updates/sec

Title:       US-09-701-586B-2
Perfect score: 2998
Sequence:    1 MAARRRSTGGGRARALNES.........PNQVRMRYLLKVQFNFLQLW 570

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:    283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters:     283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_71:*
             1: pir1:*
             2: pir2:*
             3: pir3:*
             4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                            SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1176	39.2	635	2	T01311	NAD+ ADP-ribosyltr
2	1149.5	38.3	653	2	T03656	probable NAD+ ADP-
3	1022	34.1	1011	1	JH0581	NAD+ ADP-ribosyltr
4	1017	33.9	1013	1	S04200	NAD+ ADP-ribosyltr
5	1012.5	33.8	1014	1	A29725	NAD+ ADP-ribosyltr
6	1009	33.7	1016	1	JS0428	NAD+ ADP-ribosyltr
7	1006.5	33.6	500	2	S26057	NAD+ ADP-ribosyltr
8	984.5	32.8	998	2	S31735	NAD+ ADP-ribosyltr
9	979.5	32.7	994	1	A47474	NAD+ ADP-ribosyltr
10	966	32.2	983	2	T51353	NAD+ ADP-ribosyltr
11	966	32.2	1009	2	C84719	probable poly (ADP
12	964	32.2	996	1	S42208	NAD+ ADP-ribosyltr
13	961.5	32.1	969	2	T03657	NAD+ ADP-ribosyltr
14	685.5	22.9	727	2	T18600	hypothetical prote
15	674.5	22.5	459	2	T08713	NAD+ ADP-ribosyltr
16	526	17.5	538	2	T20414	hypothetical prote
17	331.5	11.1	135	2	PN0494	NAD+ ADP-ribosyltr
18	165.5	5.5	2004	2	D88948	protein ZK1005.1
19	141.5	4.7	181	2	T03058	NAD+ ADP-ribosyltr
20	134.5	4.5	1170	2	A72287	hypothetical prote
21	133	4.4	805	2	S48411	SEC6 protein - yea
22	130.5	4.4	1005	2	A64465	hypothetical prote
23	126	4.2	880	2	F75103	conserved hypothet
24	126	4.2	2058	2	A59267	myosin X - human
25	125	4.2	1051	2	T18351	lmp1 protein - Myc
26	125	4.2	1365	2	T30822	lmp1 protein - Myc
27	122.5	4.1	2401	2	T28676	rhoptry protein -
28	122	4.1	881	2	I84737	kinesin heavy chai
29	121	4.0	631	2	JC4298	hyaluronan recepto
30	120.5	4.0	1270	2	T22615	hypothetical prote
31	120	4.0	1119	2	B70126	surface-located me
32	120	4.0	3259	1	A56539	giantin - human
33	119	4.0	1199	2	T29145	hypothetical prote
34	119	4.0	1780	2	T17272	hypothetical prote
35	118.5	4.0	1325	2	T42722	male-enhanced anti
36	117.5	3.9	1940	2	A29320	myosin heavy chain
37	117.5	3.9	2231	2	D71870	hypothetical prote
38	117	3.9	963	1	A41919	kinesin heavy chai
39	116.5	3.9	725	1	JC5016	hyaluronan recepto
40	116.5	3.9	886	2	H69378	conserved hypothet
41	116.5	3.9	1827	2	T16270	hypothetical prote.
42	116	3.9	504	2	S28298	hypothetical prote
43	116	3.9	578	2	T40984	transcription fact
44	115.5	3.9	540	2	S38085	hypothetical prote
45	115.5	3.9	600	2	A33658	paraflagellar rod

```
                            ALIGNMENTS

RESULT   1
T01311
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana
N:Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C:Accession: T01311; S65662
R:Kalicki, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana T14P8.
A:Reference number: Z14290
A:Accession: T01311
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-635 <KAL>
A:Cross-references: EMBL:AF069298; NID:g3193282; PID:g3193299
A:Experimental source: cultivar Columbia
R:Lepiniec,L.; Babiychuk, E.; Kushnir, S.; Van Montagu, M.; Inze, D.
FEBS Lett. 364, 103-108, 1995
A:Title: Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(AD
A:Reference number: S65662; MUID:95269779
A:Accession: S65662
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-115; /GT; /116-635 <LEP>
A:Cross-references: EMBL:248243; NID:g853721; PIDN:CAA88288.1; PID:g853722
C:Genetics:
A:Gene: PARP
A:Map position: 4
A:Introns: 28/3; 116/1; 162/3; 190/1; 268/3; 312/1; 330/3; 355/2; 371/1; 387/3; 408/3
A:Note: T14P8.19
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger

Query Match             39.2%;   Score 1176;  DB 2;   Length 635;
Best Local Similarity   43.9%;   Pred. No. 9.9e-67;
Matches  254;   Conservative  98;  Mismatches 187;   Indels  40;  Gaps  11;

Qy      2  AARRRRSTGGGR----ARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKA  57
           |||::|             :|| |||||||::|||:.::::::||:::|::||:||||||   :|
Db     81  AIKRGLDTTGTKKDLLERLCNDANNVSN--APVKSS------------NDEA  118

Qy     58  NKDRT---EDKQDESVKALLKGKAPVDPECTAKV-GKAHVYCEGNDVYDVMLNQTNLQF  113
           :|:|:|    |:|::|    :: :||||||| ||||:::|||||::|||||:|:|||::
Db    119  EDDNNGEEEKKDEKIVTATKGAAVLDQWIPDEIKSQYHVLQRGDDVYDAILNQTNVRD  178

Qy    114  NNNKYYILQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKN  173
           |||:|||||::|| :|:: :: |||||| :: ||||:|     || ||||| :|:||||||
Db    179  NNNKFFVLQVLESDSKKTYMVTTRWGRVGVKGQSKLDGPYDSWDRAIEIFTNKFNDKTKN  238

Qy    174  NWEDREKFEKVPGKYDMLQMDYATNTQDEEEKKKEESLKSPLKPE-SQLDLRVQELIKLI  232
           | ||:|     ||||   : :|||  :|||:  ::|||  :|||| :||| :::|| |||
```

```
Db  239  YWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNNDIPSSSSEVKPEQSKLDTRVAKFISLI  298
Qy  233  CNVQAEMEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEF   292
Db  299  CNVSMAQHMMEIGYNANKLPLGKISKSTISKGYEVLKRISEVIDRYDRTR-LEELSGEF   357
Qy  293  YTRIPHDFGLR--TPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYR  350
Db  358  YIYVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKILLSVDPGLQDDPLYYHYQ 417
Qy  351  NLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE--DLH   408
Db  418  QLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVEIAQLFRASRAVEADRFQQFSSSK   477
Qy  409  NRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKN   468
Db  478  NRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKGVYFADMFSKSANYCYANTGAN   537
Qy  469  TGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTL-NGSTVPLGP   527
Db  538  DGVLLLCEVALGDMNELLYSDYNADNLPPGKLSTKGVGKTAPNPSEAQTLEDGVVVPLGK   597
Qy  528  ASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNF  566
Db  598  PVERSCSKG---MLLYNEYIVYNVEQIKMRYVIQVKFNY  633
```

RESULT   2
T03656
probable NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Oct-1999
C;Accession: T03656
R;Babyichuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Monta
submitted to the EMBL Data Library, November 1997
A;Description: Higher plants possess two poly(ADP-ribose) polymerases.
A;Reference number: Z14991
A;Accession: Z14991
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-653 <BAB>
A;Cross-references: EMBL:AJ222588; NID:e1264090; PIDN:CAA10888.1; PID:e1264091
C;Genetics:
A;Gene: PARP
C;Keywords: DNA binding; glycosyltransferase; NAD: pentosyltransferase

Query Match   38.3%;   Score 1149.5; DB 2; Length 653;
Best Local Similarity   44.0%;   Pred. No. 4.9e-65;
Matches 255; Conservative 89; Mismatches 191; Indels 45; Gaps 13;

```
Qy    1  MAARRRSTGGGRARALNESKRVVNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD   60
Db  102  LAKARGVAANGGKKDVIQRLLSATAGPAAVADGGPL--------GAKEVIKGG---DE  148
Qy   61  RTEDKQDESVKALLLKGKAPVDPECTAKVG-KAHVYCEGNDVYDVMLNQTNLQFNNNKYY  119
Db  149  EVEVKKEKMVTA-TKKGAAVLDQHIPDHIKVNYHVLQVGDEITDATLNQTNVGDNNNKFY  207
Qy  120  LIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE  179
Db  208  IIQVLESDAGGSPMVYNRWGRVGVRGQDKLHGPSTRDQAIYEFEGKFHNKTNNHWSDRK   267
Qy  180  KFEKVPGKYDMLQMDYATNTQDEEETKKE---ESLKSPLKPESOLDLRVQELIKLICNVQ  236
Db  268  NFKCYAKKYTWLEMDYG------ETEKEIEKGSITDQIK-ETKLETRIAQFISLICNIS   319
Qy  237  AMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRI  296
Db  320  MMKQRMVEIGYNAEKLPLGKLRKRATILKGYHVLKRISDVISKADR-RHLEQLTGEFYTVI 378
Qy  297  PHDFGLRTPP--LIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHC  354
```

```
Db  379  PHDFGFRKMREFIIDTPQKLKAKLEMVEALGEIEIATKLLEDDSSDQDDPLYARYKQLHC  438
Qy  355  ALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE--DLHNRML  412
Db  439  DFTPLEADSDEYSMKSYLRNTHGKTHSGYTVDIVQIFKVSRHGETERFQKFASTRNRML   498
Qy  413  LWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLL  472
Db  499  LWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGKGVYFADMFSKSANYCYASEACRSGVL  558
Qy  473  LLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHF-VTLNGSTVPLG-----  526
Db  559  LLCEVALGDMNELLNADYDANNLPKGKLRSKGVGQTAPNMVESKVADDGVVVPLGEPKQE  618
Qy  527  PASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNF  566
Db  619  PSKRGGLL------YNEYIVYNVDQIRMRYVLHVNFNF  650
```

RESULT   3
JH0581
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - chicken
N;Alternate names: poly(ADP-ribose) synthase
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JH0581
R;Ittel, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.
Gene 102, 157-164, 1991
A;Title: Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence an
A;Reference number: JH0581; MUID:91340148
A;Accession: JH0581
A;Molecule type: mRNA
A;Residues: 1-1011 <ITT>
A;Cross-references: EMBL:X52690; NID:g63742; PIDN:CAA36917.1; PID:g63743
C;Comment: This protein is a chromatin-bound enzyme.
C;Comment: This enzyme catalyzes DNA-dependent post-translational modifications of va
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD: nucleus; pentosyltransferase; zinc

Query Match   34.1%;   Score 1022; DB 1; Length 1011;
Best Local Similarity   40.6%;   Pred. No. 1e-56;
Matches 217; Conservative 110; Mismatches 187; Indels 20; Gaps 11;

```
Qy   41  RCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGND  100
Db  494  KCSKPANMK---SAGKVKEEQGPSKSEKKMK-LTVKGGAAVDPD-SGLEDSAHVFEKGGK  548
Qy  101  VYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAK  160
Db  549  IFSATGLVDIVKGTNSYYKLQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAV   608
Qy  161  EIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ  220
Db  609  EHFLNLYEEKTGNSWHSK-NFTKYPKKFYPLEIDYG--QDEEAVRK---LTVSAGTKSK   661
Qy  221  LDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQ  280
Db  662  LAKPIQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILNEVQQAVSDGG  721
Qy  281  HGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIATKLIVK-TELQ 339
Db  722  SESQIDLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRGGNED   781
Qy  340  SPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE  399
Db  782  GDKDPIDINYEKLRTDIKVVDKSDEEAKIIKQYVKNTHAATHNAYDLKVVEIFIREGE   841
Qy  400  KEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKS  457
Db  842  SQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGMFGKGIYFRDMVSKS   901
Qy  458  ANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVT  517
```

```
Db    902 ANYCHTSQADPIGLILLGEVALGNMYELKNASHITK-LPKGKHSVKGLGKTAPDPTATTT 960
          ||||      ::  |||||| |||||  ||   :||||| ||||| :| |
Qy    518 LNGSTVPLGPASDTGILNPDGYTLNYEYIVYNPNQVRMRYLLKVQFNF-LQLW 570
          |:| |||| |||   : | ||||||| ||  :||:||::::|| :||
Db    961 LDGVEVPLGNGISTGI---NDTCLLYNEYIVYDVAQVNLKYLLKLKFNYKTSLW 1011

RESULT  4
S04200
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S04200
R:Huppi, K.; Bhatia, K.; Siwarski, D.; Klinman, D.; Cherney, B.; Smulson, M.
Nucleic Acids Res. 17, 3387-3401, 1989
A:Title: Sequence and organization of the mouse poly (ADP-ribose) polymerase gene.
A:Reference number: S04200; MUID:89263780
A:Accession: S04200
A:Molecule type: DNA
A:Residues: 1-1013 <HUP>
A:Cross-references: EMBL:X14206; NID:g49893; PIDN:CAA32421.1; PID:g49894
C:Genetics:
A:Map position: 1
C:Superfamily: NAD+ ADP-ribosyltransferase
C:Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc fi

Query Match           33.9%;  Score 1017;  DB 1;  Length 1013;
Best Local Similarity 40.1%;  Pred. No. 2.1e-56;
Matches   219;  Conservative  113;  Mismatches  178;  Indels  36;  Gaps  12;

Qy     31 EDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVG 90
          ::||:|:||    |  :: |||| |  |||| | | :| ||||
Db        KSAAPSKKSKGCFKEEG-------------VNKSEKRMK-LTLKGGAAVDPD-SGLEH 540

Qy     91 KAHVYCEGNDVYDVMLNQTNLQFNNNKYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLV 150
          |||   ||| |  ||| | :  || |  :|||||| :: :
Db    541 SAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRLGTVIGSNKL 600

Qy    151 ACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEETKKEES 210
          | |  :| |   |||: | :::||   |||| |  |  |    :|   |
Db    601 EQMPSKEEAVEQFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYG---QDEEAVK---- 652

Qy    211 LKSPLKP--ESQLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPIGKLTVAQIKAGYQS 268
          : ::|    ::| ||| |: | :  : : :  ||    |::| | :|||||:
Db    653 -KLTVKPGTKSKLPKPVQELVGMIFDVDSMKKALVEYEIDLQKMPLGKLSRRQIQAAYSI 711

Qy    269 LKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIE 328
          |  :       | | |  : |  |||  || || |   :|||  |: |  | | |:||| |
Db    712 LSEVQQPVSQGSSESQIIDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIE 771

Qy    329 IAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMT 387
          : ||||:  ||  || |:  | :|: | |||||| |  :  | :| :|| ||:| ||
Db    772 VAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDLE 831

Qy    388 LLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFG 445
          ::::|:|:| ||||   |||||||| || ||||:| ||: |||  :|| | ||||| |||
Db    832 VIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFG 891

Qy    446 KGIYFADMSSKSANYCFASRLKNTGLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGL 505
          ||||||||| |||||| |:    |:|||| |: | | |||  | | ::: ||||| |||
Db    892 KGIYFADMVSKSANYCHTSQGDPIGLIMLGEVALGNMYELKHASHISK-LPKGKHSVKGL 950

Qy    506 GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYEYIVYNPNQVRMRYLLKVQFN 565
          || || ||| |||||| |||| |||||||||||||||||||||:|| |: ||||:||||
Db    951 GKTPDPSASITLEGVEVPLGTGIPSGV---NDTALLYNEYIVYDIAQVNLKYLLKLKFN 1007

Qy    566 F-LQLW 570
          | |||
Db   1008 FKTSLW 1013
```

```
RESULT  5
A29725
NAD+ ADP-ribosyltransferase (EC 2.4.2.30), nuclear - human
N:Alternate names: poly (ADP-ribose) polymerase; poly (ADP-ribose) synthetase; poly(A
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29725; A28498; A39976; A26901; I38096; B33321; A33321; A35635; A61559;
R:Uchida, K.; Morita, T.; Sato, T.; Ogura, T.; Yamashita, R.; Noguchi, S.; Suzuki, H.
Biochem. Biophys. Res. Commun. 148, 617-622, 1987
A:Title: Nucleotide sequence of a full-length cDNA for human fibroblast poly(ADP-ribo
A:Reference number: A29725; MUID:88076933
A:Accession: A29725
A:Molecule type: mRNA
A:Residues: 1-69,'Q',71-1014 <UCH>
A:Cross-references: GB:M18112; NID:g190166; PIDN:AAA60137.1; PID:g190167
R:Kurosaki, T.; Ushiro, H.; Mitsuuchi, Y.; Suzuki, S.; Matsuda, M.; Katu
J. Biol. Chem. 262, 15990-15997, 1987
A:Title: Primary structure of human poly (ADP-ribose) synthetase as deduced from cDNA
A:Reference number: A28498; MUID:88058958
A:Accession: A28498
A:Molecule type: mRNA
A:Residues: 1-16,'E',18-211,'K',213-236,'R',238-366,'H',369-1014 <KUR>
A:Cross-references: GB:J03473
R:Cherney, B.W.; McBride, O.W.; Chen, D.; Alkhatib, H.; Bhatia, K.; Hensley, P.; Smul
Proc. Natl. Acad. Sci. U.S.A. 84, 8370-8374, 1987
A:Title: cDNA sequence, protein structure, and chromosomal location of the human gene
A:Reference number: A39976; MUID:88068596
A:Accession: A39976
A:Molecule type: mRNA
A:Residues: 1-49,'D',51-612,'Q',614-907,'Y',909-939,'R',941-979,'I',981-1014 <CHE>
A:Cross-references: GB:J03030
A:Note: the authors translated the codon ATA for residue 980 as Asn
R:Suzuki, H.; Uchida, K.; Shima, H.; Sato, T.; Okamoto, T.; Kimura, T.; Miwa, M.
Biochem. Biophys. Res. Commun. 146, 403-409, 1987
A:Title: Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and express1
A:Reference number: A26901; MUID:87298455
A:Accession: A26901
A:Molecule type: mRNA
A:Residues: 441-610,'N',612-880:921-1014 <SU2>
A:Note: the sequence figure has an omission of forty residues
R:Ogura, T.; Nyunoya, H.; Takahashi-Masutani, M.; Miwa, M.; Sugimura, T.; Esumi, H.
Biochem. Biophys. Res. Commun. 167, 701-710, 1990
A:Title: Characterization of a putative promoter region of the human poly(ADP-ribose)
A:Reference number: I38096; MUID:90211250
A:Accession: I38096
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-40 <RES>
A:Cross-references: EMBL:X16674; NID:g510112; PIDN:CAA34663.1; PID:g1017423
R:Auer, B.; Nagl, U.; Herzog, H.; Schneider, R.; Schweiger, M.
DNA 8, 575-580, 1989
A:Title: Human nuclear NAD(+) ADP-ribosyltransferase(polymerizing): organization of t
A:Reference number: A33321; MUID:90091744
A:Accession: B33321
A:Molecule type: DNA
A:Residues: 38-43;93-98;132-137;204-209;237-242;276-281;335-340;384-389;431-436;512-5
A:Cross-references: GB:M22953
A:Note: the authors translated the codon GTG for residue 54 as Glu
A:Note: these fragments represent intron-exon boundaries
A:Accession: A33321
A:Molecule type: DNA
A:Residues: 38-43;93-98;132-137;204-209;237-242;276-281;335-340;384-389;431-436;512-5
A:Cross-references: GB:M22953
A:Note: the authors translated the codon GTG for residue 54 as Glu
A:Note: these fragments represent intron-exon boundaries
A:Accession: A33321
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 16-66;96;121-159,'D',161-167 <AU2>
A:Note: these fragments represent a zinc finger-containing DNA-binding region
R:Gradwohl, G.; Menissier de Murcia, J.; Molinete, M.; Simonin, F.; Koken, M.; Hoeijm
Proc. Natl. Acad. Sci. U.S.A. 87, 2990-2994, 1990
A:Title: The second zinc-finger domain of poly(ADP-ribose) polymerase determines spec
A:Reference number: A35635; MUID:90222155
A:Accession: A35635
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 12-26,'T',28-66;116-166 <GRA>
R:Schneider, R.; Auer, B.; Kuehne, C.; Herzog, H.; Klocker, H.; Burtscher, H.J.; Hirs
```

Eur. J. Cell Biol. 44, 302-307, 1987
A;Title: Isolation of a cDNA clone for human NAD (+): protein ADP-ribosyltransferase.
A;Reference number: A61559
A;Accession: A61559
A;Molecule type: mRNA
A;Residues: 381-420:682-710 <SCH>
R;Yokoyama, Y.; Kawanoto, T.; Mitsuuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Terashi
Eur. J. Biochem. 194, 521-526, 1990
A;Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.
A;Reference number: S14010; MUID:91099327
A;Accession: S14010
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-95 <YOK>
C;Genetics:
A;Gene: GDB:ADPRT; PPOL
A;Cross-references: EMBL:X56140; NID:g35286; PIDN:CAA39606.1; PID:g825702
A;Cross-references: GDB:119508; OMIM:173870
A;Map position: 1q41-1q42
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransfer

```
Query Match              33.8%;   Score 1012.5;  DB 1;  Length 1014;
Best Local Similarity    40.5%;   Pred. No. 4.1e-56;
Matches   221;  Conservative 111;  Mismatches 185;  Indels  29;  Gaps  11;

Qy    29  APEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAK   88
Db   494  APRGKSGAALSKKSKCQV----------KEEGINKSEKRMK-LTLKGGAAVDPD-SGL    539

Qy    89  VGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHS  148
Db   540  EHSAHVLEKGGKVFSATLGLVDIVKGTNSYKLQLLEDDKENRVWIFRSWGRVGTVIGSN   599

Qy   149  LVACSGNLNKAKEIFQKFLDKTKNNWEDREKFEKVPGKYDMLQMDYAINTQDEEETKE    208
Db   600  KLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYG--QDEEAVKK-  654

Qy   209  ESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQS  268
Db   655  --LTVNPGTKSKLPKPVQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSI  712

Qy   269  LKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIE  328
Db   713  LSEVQQAVSQGSSDSQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIE  772

Qy   329  IAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMT  387
Db   773  VAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATTHNAYDLE  832

Qy   388  LLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFG  445
Db   833  VIDIFKIEREGEBCQRYKPFKQLHNRRLLWHGSRTNFAGIISQGLRIAPPEAPVTGYMFG  892

Qy   446  KGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGL  505
Db   893  KGIYFADMVSKSANYCHTSQGDPIGLILGEVALGNMYELKHASHISK-LPKGKHSVKGL   951

Qy   506  GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNEYIVYNPNQVRMRYLLKVQFN    565
Db   952  GKTTPDPSANISLDGVDVPLGTGISSGV---NDTSLLYNEYIVYDIAQVNLKYLLKLKFN 1008

Qy   566  F-LQLW   570
Db  1009  FKTSLW  1014
```

RESULT   6
JS0428
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine
N;Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate rib

C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: JS0428; S00328; A30458
R;Saito,I.
submitted to JIPID, February 1990
A;Reference number: JS0428
A;Accession: JS0428
A;Molecule type: mRNA
A;Residues: 1-1016 <SAI>
A;Experimental source: thymus
R;Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; T
Eur. J. Biochem. 171, 571-575, 1988
A;Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int
A;Reference number: S00328; MUID:88151954
A;Accession: S00328
A;Molecule type: mRNA
A;Residues: 648-714;838-904 <TAN>
A;Cross-references: EMBL:X06986
A;Accession: A30458
A;Molecule type: protein
A;Residues: 658-685;689-696;893-901 <TA2>
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-loo
F;21-51/Region: zinc finger
F;128-165/Region: zinc finger
F;200-220/Region: helix-turn-helix motif
F;224-231/Region: nuclear location signal
F;250-270/Region: helix-turn-helix motif
F;494-501/Region: nucleotide-binding motif A (P-loop)
F;890-903/Region: nucleotide binding #status predicted

```
Query Match              33.7%;   Score 1009;   DB 1;  Length 1016;
Best Local Similarity    40.4%;   Pred. No. 6.8e-56;
Matches   222;  Conservative 108;  Mismatches 187;  Indels  32;  Gaps  12;

Qy    26  GNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPEC   85
Db   496  GPKGKSGAAPSKKS----------KGPV-----KEEGTNKSEKRMK-LTLKGGAAVDPD-  538

Qy    86  TAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMG  145
Db   539  SGLEHNAHVLEKGGKVFSATLGLVDIVKGTNSYKLQLLEDDKESRTWIFRSWGRVGTVI   598

Qy   146  QHSLVACSGNLNKAKEIFQKFLDKTKNNWEDREKFEKVPGKYDMLQMDYAINTQDEEET   205
Db   599  GSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKHPKKFPYPLEIDYG---QDEEAV 654

Qy   206  KKEESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAG  265
Db   655  KK---LTVNPGTKSKLPKPVQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAA  711

Qy   266  YQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALG  325
Db   712  YSILSEVQQALSQGSSDSHILDLSNRFYTLIPHDFGMKKPPLLNNANSVQAKVEMLDNLL  771

Qy   326  DIEIAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDY  384
Db   772  DIEVAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDKSEEAEIIRKYVKNTHATTHNAY   831

Qy   385  TMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGY  442
Db   832  DLEVVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGY  891

Qy   443  MFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHST  502
Db   892  MFGKGIYFADMVSKSANYCHTSQGDPIGLILGEAALGNMYELKHARHISK-LPKGKHSV   950

Qy   503  KGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNEYIVYNPNQVRMRYLLKV    562
Db   951  KGLGKTTPDPSASITVDGVEVPLGTGISSGV---NDTCLLYNEYIVYDIAQVHLKYLLKL 1007

Qy   563  QFNF-LQLW  570
```

Db   1008 KFNFKTSLW 1016
          :||| ||
Qy   539 YT---LNYNEYIVYNPNQVRMRYLLKVQFNF-LQLW 570

RESULT 7
S26057
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 18-Jun-1999
C;Accession: S26057; S78453; I52331
R;Thibodeau, J.; Gradwohl, G.; Dumas, C.; Clairoux-Moreau, S.; Brunet, G.; Penning, C.;
Biochem. Cell Biol. 67, 653-660, 1989
A;Title: Cloning of rodent cDNA encoding the poly(ADP-ribose) polymerase catalytic domain
A;Reference number: I52331; MUID:90027702
A;Accession: S26057
A;Molecule type: mRNA
A;Residues: 1-500 <THI>
A;Cross-references: EMBL:X65497; NID:g56849; PIDN:CAA46478.1; PID:g56850
R;Potvin, F.
submitted to the EMBL Data Library, March 1992
A;Reference number: S78453
A;Accession: S78453
A;Molecule type: mRNA
A;Residues: 1-124,'H',126-127,'A',129-238,'D',240-500 <POT>
A;Cross-references: EMBL:X65497
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransfer

Query Match   33.6%;   Score 1006.5;   DB 2;   Length 500;
Best Local Similarity 42.6%;   Pred. No. 3.7e-56;
Matches 220;   Conservative 101;   Mismatches 168;   Indels 27;   Gaps 13;

Qy   64 DKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQL 123
          :|| :| ||||| |||| ||| ||| |||| :|||:|||||||||| |||||||
Db   3 NKSEKRMK-LTLKGGAAVDPD-SGLEHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQL 60

Qy   124 LEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEK 183
          |||| |||||| ||||||||||||:||||:|||||||||||||||| :|| ||||||
Db   61 LESDKESRYWIFRSWGRVGTVIGSNKLEQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTK 119

Qy   184 VPGKYDMLQMDYATNTQDEEETKEEESLKSPLKP--ESQLDLRVQELIKLICNVQAMEEM 241
          | | || || | |::| :|::| :|:||:|:||:|
Db   120 YPKKFYPLEIDYG---QDEEAVK----KLAVKPGTKSKLPKPVQELVGMIFDVESMKKA 171

Qy   242 MMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFG 301
          :|:| | ||| | | ::| : | :: |::| | | |||||||||||||||| ||||
Db   172 LVEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQAVSQGSSESQILDLSNRFYTLIPHDFG 231

Qy   302 LRTPPLIRTQKELSEKIQLEALGDIEIAIKLVK-TELQSPEHPLDQHYRNLHCALRPLD 360
          ||| |:|||||||||||:|||||| ||||||| :|||:|||| | |||||||||||||
Db   232 MKKPPLLNNTDSVQAKVEMLDNLLDIEVAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVD 291

Qy   361 HESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSR 418
          |:: |||| :|:||| :|:|||| ||||||||||||||||| ||||||||||||||
Db   292 RDSEAEVIRKYVKNTHATTHNAYDLEVIDIFKIEREGESQRYKPFRQLHNRRLLWHGSR 351

Qy   419 MSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVA 478
          :| ||||||||||||||||:||||||||||||||||||||||| |||||| :||| ||
Db   352 TTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANYCHTSQGDPIGLILLGEVA 411

Qy   479 LGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDG 538
          | | |||||||||||||||||||||||||| ||||||||||||:|||| :|:|||||||
Db   412 LGNMYELKHASHISK-LPKGKHSVKGLGKTAPDPSASITDLGVEVPLG----TGI--PSG 464

Qy   539 YT---LNYNEYIVYNPNQVRMRYLLKVQFNF-LQLW 570
          |||||| || :|||| | :|||||| |||||||
Db   465 VNDTCLLYNEYIVDIAQVNLKYLLKLKFNFKTSLW 500

RESULT 8
S31735
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - African clawed frog (fragment)

N;Alternate names: poly ADP-ribose polymerase
C;Species: Xenopus laevis (African clawed frog)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S31735; PN0495
R;Saulier-le Drean, B.M.
submitted to the EMBL Data Library, May 1992
A;Accession: S31735
A;Molecule type: mRNA
A;Residues: 1-998 <SAU>
A;Cross-references: EMBL:Z12139; NID:g64967; PIDN:CAA78126.1; PIDN:g1334661
R;Ozawa, Y.; Uchida, K.; Uchida, M.; Ami, Y.; Kushida, S.; Okada, N.; Miwa, M.
Biochem. Biophys. Res. Commun. 193, 119-125, 1993
A;Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)polymera
A;Reference number: PN0494; MUID:93277538
A;Accession: PN0495
A;Molecule type: mRNA
A;Residues: 742-745,'E',747-876 <OZA>
C;Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and diff
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pento

Query Match   32.8%;   Score 984.5;   DB 2;   Length 998;
Best Local Similarity 39.1%;   Pred. No. 2.4e-54;
Matches 222;   Conservative 109;   Mismatches 196;   Indels 41;   Gaps 14;

Qy   21 KRVNNGNTAPE------------DSSPAKKTRRCQRQESKKMPVAG------GKANKDRTEDKQD 67
          | :| | : :|: | ::|:|| || ||
Db   444 KEVESGKSVQELLSQFGISSWGAEIKQEAVQPTEKQPSSGPVAGKSSGKVKEEKGSNKSE 503

Qy   68 ESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDD 127
          || || || || |: || | || | ||||||| |||||||
Db   504 KKMK-LTVKGGAAIDPDSELE-DSCHVLETGGKIFSATLGLVDITRGTNSYYKLQLIEHD 561

Qy   128 AQRNFSVWMRWGRVGK-MGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPG 186
          ||||||| |||||| ||||:||||:||||||||||||||||| :|| |||||| |
Db   562 RDSRYWFRSWGRVGTVIGSKKLEEMSSK-EDAIEHFLNLYQDKTGNAWHS-PNFTKYPK 619

Qy   187 KYDMLQMDYATNTQDEEETKEEESLKSPLKPESQLDRVQELIKLICNVQAMEEMMEMK 246
          | || | :|:| ||::| :|:| :|||:| ||||||| ||||||:|||| ::
Db   620 KFYPLEIDYG---QEEDVKK----LSVGAGYKSKLAKPVQELIKLIFDVESMKKAMVEFE 673

Qy   247 YNTKRAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPP 306
          |||| |||||| :|:| : | :: |::|:| | |||||||||||||||| ||||||||
Db   674 IDLQKMPLGKLSKRQIQSAYSILSQVQQAVSESLSEARLLDLSNQFYTLIPHDFGMKKPP 733

Qy   307 LIRTQKELSEKIQLEALGDIEIAIKLVKTELQSPE-HPLDQHYRNLHCALRPLDHESYE 365
          || ::|||||||::|||| ||:|||||||| | | | |||||||| |||| :|
Db   734 LLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDGEKDPIDVKYEIKTDIKVVAKDSEE 793

Qy   366 FKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWV 423
          |:::|:||| :|:|||| |||||||||||||||||| ||||||||||||||||:
Db   794 SRIICDYVKNTHADTHNAYDLEVLEIFKIDREGEYQRYKPFKQLHNRQLLWHGSRTTNFA 853

Qy   424 GILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCN 483
          |||| |||||||||||||||||||||||||||||||| |||||| :||| :|| :|
Db   854 GILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANYCHAMPGSPIGLILLGEVALGNMH 913

Qy   484 ELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPA----SDTGILNPDG 538
          ||||| ||||||||||||||||| ||| ||||||||| ||| | :|:|||||||
Db   914 E-LKAASQITKLPKGKHSVKGLGRTAPDPSATVQLDGVDVPLGKGTSANISDTSLL---- 968

Qy   539 YTLNYNEYIVYNPNQVRMRYLLKVQFNF 566
          ||||| || :|||| | :||||||||||
Db   969 ---YNEYIVDIAQVNLKYLLKLKFNY 992

RESULT 9
A47474
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C;Accession: A47474
R;Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.
Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993
A;Title: Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase: leucine zipper
A;Reference number: A47474; MUID:93234521
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-994 <UCH>
A;Cross-references: GB:D13806; GB:D13807; GB:D13808; NID:g303545; PIDN:BAA02964.1; PID:g
A;Note: sequence extracted from NCBI backbone (NCBIN:129703, NCBIP:129704)
C;Genetics:
A;Gene: FlyBase:Parp
A;Cross-references: FlyBase:FBgn0010247
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase

```
Query Match       32.7%;   Score 979.5;  DB 1;   Length 994;
Best Local Similarity 38.3%;  Pred. No. 4.9e-54;
Matches 215;  Conservative 110;  Mismatches 199;  Indels 37;  Gaps 11;

Qy   12 GRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQE--SKKMPVAGGKANKDRTEDKQDES 69
           : :::  :      ::       :       : :   || |  || : |||||||||||
Db  463 GAIKYIN-STCICSWGTDPKSRIPKETTKSLNSNSIYTKSMPVS------------ 505

Qy   70 VKALLLGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYLIQLLEDDAQ 129
           :  ::   |||   |  ||||||||||  : ||  :  : |  :||| ||||||||||||
Db  506 -RTFKVKDGLAVDPD-SGLEDIAHVYDSNNKYSVVLGLTDIQRNKNSYYKVQLLKADKK 563

Qy  130 RNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYD 189
            : |  | |||:|   | | : :||| |||| |: :||||||:  :|| ||||| |||||
Db  564 EKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVKRTGRMY 623

Qy  190 MLQMDYATNTQDEEETKKEESLKSPLKPESQLDRVQELIKLICNVQAMEEMMEMKYNT 249
           :: :   : :      |:: :     || : |:||||||||||||||||||||||||| :
Db  624 PIEIQY----DDDQKLVKHESHFF----TSKLEISVQNLIKLIFDIDSMNKTLMEFHIDM 675

Qy  250 KKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFEGLRTPPLIR 309
           :|||||| : :|| ||| |:  :|  | || :    :    | ||| ||| |  :  |:|
Db  676 DKMPLGKLSAHQIQSAYRVVKEIYNVLECGSNTAKLIDATNRFYTLIPHNFGVQLPTLIE 735

Qy  310 TQKELSEKIQLLEALGDIEIAIKLVKTE-LQSPEHPLDQHYRNLHCALRPLDHESYEFKV 368
            :: |  :: :|  :| |||| : : ||  : ||  ||  || || || |:  |: ||:
Db  736 THQOIEDLRQMLDSLAEIEVAYSIIKSEDVSDACNPLDNHYAQIKTQLVALDKNSEEFSI 795

Qy  369 ISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGIL 426
           :||| :||||:||:|  : :  : || :::| |||  |||||||:||||||||:|:|||
Db  796 LSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARRFKPFKKLHNRKLLWHGSRLTNFVGIL 855

Qy  427 SHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELL 486
           |||||||||||| ||||||||||||| ||||||| | : :|| | ::|||||||  :   :
Db  856 SHGLRIAPPEAPPTGYMFGKGIYFADMVSKSANYCCTSQQNSTGLMLLSEVALGDMMECT 915

Qy  487 EANPKAEGLLQGKHSTKGLGKMAPS-SAHFVTLNGSTVPLGPASDTGILNPDGYTLNVNE 545
            :   : | | ||||| | | | |  |||| :   | ||  :|| || ||: |  ||::|
Db  916 SAK-YINKLSNNKHSCFGRGRTMPDDPTKSYIRSDGVEIPYGETITDEHLKS---SLLYNE 971

Qy  546 YIVYNPNQVRMRYLLKVQFNF 566
           |||: :  | ||||||| |
Db  972 YIVYDVAQVNIQLFRMEFKY 992
```

RESULT 10
T51353
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) [imported] - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C;Accession: T51353
R;Doucet-chabeaud, G.; Kazmaier, M.
submitted to the EMBL Data Library, December 1998
A;Reference number: 225379

A;Accession: T51353
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-983 <DOU>
A;Cross-references: EMBL:AJ131705; PIDN:CAA10482.1
A;Experimental source: cultivar landsberg erecta
C;Genetics:
C;Gene: parp-1
C;Function:
A;Description: ADP-ribose polymer synthesis
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

```
Query Match       32.2%;   Score 966;  DB 2;   Length 983;
Best Local Similarity 38.8%;  Pred. No. 3.4e-53;
Matches 213;  Conservative 115;  Mismatches 151;  Indels 70;  Gaps 17;

Qy   45 QESKKMPVAGGKANKDRTEDKQDESVKALLLGKAPVDPECTAKVGKAHVYCEGNDVYDV 104
           ::  ::||      ::   :: ::   :| ::|:  :|    || |  :| : :|  :
Db  472 KKQRKLPF-----DKYKIED-TSESLIVTVKVKGRSAVH-EASGLQEHCHILEDGNSIYNT 524

Qy  105 MLNQTNLQFNNNKYLIQLLEDDAQRNFSVWMRWGRVG--KMGQHSLVACSGNLNKAKEI 162
           :|: :: : :|  ||| :|:||:  :    :| :||||   :| : :     : ::: |
Db  525 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGG--------NKVEEM 574

Qy  163 --------FQKKFLDKTKNNWEDREK--FEKVPGKYDMLQMDYATNTQDEEETKKEESL 211
                    : :::| ||  ||   |  |:| ||| | :  : : | : :::    :||
Db  575 SKSDAVHEFKRLFLEKTGNTWESEWEQKTNFQKQPGKFLPLDIDYGVN-------KQVAK 626

Qy  212 KSPLKPESQLDRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKK 271
           |   : ::| : : |::|:: : || :   :: : :   :||||| :  || :: |
Db  627 KEPFQTSSNLAPSLIELMKMLFDVETYRSAMEFFEINMSEMPLGKLSKHNIQKGFEALTE 686

Qy  272 IEDCI-----RAGQHGRALMEACNEFYTRIP-HDFGLRTPPLIRTQKELSEKIQLLEAL 324
           |  |:     ||   :  :::: | ||::| :    |   |:| ::    ::: :|:||
Db  687 IQRLLTESDPQPTMKESLLVDASNRFFTMIPSIH------PHIIRDEDDFKSKVMLEAL 740

Qy  325 GDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDY 384
           |||||: ::| | |   | ||: || || |  | |||: || |: :|| :||||||: :
Db  741 QDIEIASRIVGFDVDSTE-SLDDKYKKLHCDISPLPHDSEDYRLIEKYLNTTHAPTHTEW 799

Qy  385 TMTLLDLFEVEKDGEKEAF--REDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITG 441
            : ||:: ||: :| :|    |:::  ::|||||||:|::||||::|||||||||||:||
Db  800 SLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAPATG 859

Qy  442 YMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEAN----PKAEGLL 496
           |||||||||||: ||||:||   ::| | ||:|||||||  :  ::      ||   :
Db  860 YMFGKGIYFADLVSKSAQYCYTCKKNPVGLMLLSEVALGBIEHELTKAKYMDKPPR---- 914

Qy  497 QGKHSTKGLGKMAPSSAAHFVTLNGS-TVPLGPASDTGILNPDGYTLNYNEYIVVNPNQVR 555
           ||||||||||| :  |:  |:|: |  |||  || ||:    :  |:|||||||::  ||
Db  915 -GKHSTKGLGKKVPQDSEFAKWRGDVTVPCGKPVSSKVVASE--LMYNEYIVYDTAQVK 970

Qy  556 MRYLLKVQF 564
           : |||||:|
Db  971 LQFLLKVRF 979
```

RESULT 11
C84719
probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
A;Accession: C84719
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083387
A;Accession: C84719
A;Status: preliminary

```
A;Molecule type: DNA
A;Residues: 1-1009 <STO>
A;Cross-references: GB:AE002093; NID:g4433827; PIDN:AAD20677.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g31320
A;Map position: 2
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match        32.2%;    Score 966;  DB 2;  Length 1009;
Best Local Similarity  38.8%;    Pred. No. 3.5e-53;
Matches  213;  Conservative 115;  Mismatches 151;  Indels 70;  Gaps  17;

Qy   45  QESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVCEGNDVYDV 104
             :: :||:| ||  ||||||||||||||:||: |: :||  ||::|
Db  498  KKQRKLPF------DKYKIED-TSESLVTVKVKGRSAVH-EASGLQEHCHILEDGNSIYNT 550

Qy  105  MLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVG--KMGQHSLVACSGNLNKAKEI 162
             |:|||| :|  |::|::::|| : :|||||:|  | :||
Db  551  TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGG--------NKVEEM 600

Qy  163  --------FQKFLDKTKNNWEDREK---FEKVPGKYDMLQMDYATNTQDEEETKKEESL 211
                 :||| |||||| || |: ||| ||| |:| |:
Db  601  SKSDAVHEFKRLFLEKTGNTWESWEQKTNFQKQPGKFLPLDIDYGVN--------KQVAK 652

Qy  212  KSPLKPESQLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKK 271
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  653  KEPFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTE 712

Qy  272  IEDCI-----RAGQHGRALMEACNEFYTRIP--HDFGLRTPPLIRTQKELSEKIQLLEAL 324
             :|||       |||| ||| || |||:|    |   || | || |:|||:||:|||||
Db  713  IQRLLTESDPQPTMKESLLVDASNRFFTMIPSIH------PHIIRDEDDFFKSKVKMLEAL 766

Qy  325  GDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDY 384
             ||||:|| : | |  || |  ||||:||:| |::|  |  :|::| ||:||||||||
Db  767  QDIEIASRIVGFDVDSTE-SLDDKYYKLHCDISPLPHDSEDYRLIEKYLNTTHAPTHTEW 825

Qy  385  TMTLLDLFEVEKDGEKEAF---REDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITG 441
             ||:|| |||||  || :|    ||||||| |||||||||||||||||||||||||
Db  826  SLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAPATG 885

Qy  442  YMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEAN-----PKAEGLL 496
             |||||||||||||||||||: ||: | |  :: |||||| ||::|     |
Db  886  YMFGKGIYFADLVSKSAQYCYTCKKNPVGLMLLSEVALGEIHELTKAKYMDKPPR------ 940

Qy  497  QGKHSTKGLGKMAPSSAHFVTLNGS-TVPLGPASDTGILNPDGYTLNYINEYIVYNPNQVR 555
             ||||||||||||      | || | || |||      ::|||||||||||||: ||
Db  941  -GKHSTKGLGKVPQDSEFAKWRGDVTVPCGKPVSSKVKASE---LMYNEYIVYDTAQVK 996

Qy  556  MRYLLKVQF 564
             ::|||:|:|
Db  997  LQFLLKVRF 1005


RESULT 12
S42208
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - flesh fly (Sarcophaga peregrina)
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Sarcophaga peregrina
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S42208; S71496
R;Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurat
   Eur. J. Biochem. 220, 607-614, 1994
A;Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from Sarc
A;Reference number: S42208; MUID:94170813
A;Accession: S42208
A;Molecule type: mRNA
A;Residues: 1-996 <MAS>
A;Cross-references: EMBL:D16482; NID:g473742; PIDN:BAA03943.1; PID:g538248
A;Accession: S71496
A;Molecule type: protein
A;Residues: 170-188;721-736;813-819;879-885 <MAX>
C;Superfamily: NAD+ ADP-ribosyltransferase
```

```
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger
F;1-369/Domain: DNA binding #status predicted <DNA>
F;370-507/Domain: auto-modification #status predicted <AMO>
F;508-996/Domain: NAD binding #status predicted <NAD>

Query Match        32.2%;    Score 964;  DB 1;  Length 996;
Best Local Similarity  39.9%;    Pred. No. 4.7e-53;
Matches  216;  Conservative 116;  Mismatches 177;  Indels  32;  Gaps  13;

Qy   33  SSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKA 92
                                                           LKIKDGLAVDPD-SGLEDVA 528
Db  480  TDPATRITQEESKSSKSSKSIYTKSVPKSMT----------

Qy   93  HVYCEGN-DVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVG-KMGQHSLV 150
             ||| ||   | :|| | | :|||||:::::|||||: |||:| |  ||::
Db  529  HVYVSRNKEKYNVVLGITDIQKNKNSFYKLQLLESDMKNRFWVFRSWGRIGTTIGGNKLD 588

Qy  151  ACSGNLNKAKEIFQKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEES 210
             |       :: |:|| |::  :|  ||::|   |  | ||   |:
Db  589  NFS-NLVDAIVQFKELYLEKSGNHFENRENFNVVAGRMYPIDIDYA------EDSKIDLS 641

Qy  211  LKSPLKPESQLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLK 270
               :    |::  | |:: |:| |::| :| |||||:| || |||||| ::||:| :| |:
Db  642  AEHDIK--SKLPLSVQDIIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLT 699

Qy  271  KIEDCIRAGQHGRALMEACNEFYTRIPIHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIA 330
             :|| | ||    |  ||:|:|||| || || | :|||:   :| |||:| | |:  :||
Db  700  EIYELIQGGTNAKFIDATNRFYTLIPIHNFGTQSPPLLDTTEQVEQLRQMLDSLIEIECA 759

Qy  331  IKLVKTELQSPE-HPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLL 389
             : | | :   | :|:|:||  | :| |||| :|  || ::|||: ||| ||  |  :::
Db  760  YSLLQTEDSKADINPIDKHYEQLKTKLEPLDKNSEEYILLQKYVKNTHAETHKLYDLEVV 819

Qy  390  DLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKG 447
             |:| | :  |:  :   :.|||| |||||||:|:.||||||||:|||||||:|||||||
Db  820  DIFKVARQGEARRYKPFKKLHNRRLLWHGSRLTNFAGIILSHGLKIAPPEAPVTGYMFGKG 879

Qy  448  IYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGK 507
             ||||| |||||||  :  ::||||:||||||||   |  |      |: ||||| ||:|
Db  880  IYFADMVSKSANYCCTSHHNSTGLMLLSEVALGDMMECTAAKYVTK-LPNDKHSCFGRGR 938

Qy  508  MAPS-SAHFVTLNGSTVPLG-PASDTGILNPDGYTLNYINEYIVYNPNQVRMRYLLKVQFN 565
             : |    |   |  : ||  |  ::   ::: :|:||:||| |:|   ||  ||:|||:
Db  939  TMPNPSESIIREDGVEIPLGKPITNDSLKS----SLLYNEFIIYDIAQVNIQYMLRMNFK 994

Qy  566  F 566
             |
Db  995  Y 995


RESULT 13
T03657
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) 2 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
C;Accession: T03657
R;Babiychuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Mo
   submitted to the EMBL Data Library, November 1997
A;Description: Higher plants possess two poly(ADP-ribose) polymerases.
A;Reference number: Z14992
A;Accession: T03657
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-969 <BAB>
A;Cross-references: EMBL:AJ222589; PIDN:CAA10889.1
C;Genetics:
A;Gene: PARP2
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match        32.1%;    Score 961.5;  DB 2;  Length 969;
```

```
Best Local Similarity   39.3%;   Pred. No. 6.5e-53;
Matches  223;   Conservative  104;   Mismatches  179;   Indels  61;   Gaps  18;

Qy   32  DSSPAKKTRRCQRQESKKMPVAGG------KANK-------DRTEDKQDESVKALLL  75
             ::: ||  |:|:         |:|   ||        : : :|
Db  428  ENAEVRKARRL------KIPIVREGYIGECVKKNKMLPFDLYKLENALESSKGSTVTVKV  481

Qy   76  KGKAPVDPECTAKVGKAHYCEGNDVDYDVMLNQTNLQFNNNKYLIQLLEDDAQRNFSVW  135
            ||||: ::|:::|||||||||||||||||||||||||||||||||:::|| |
Db  482  KGRSAVH-ESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYVVLQIIEQDDGSECYVF  540

Qy  136  MRWGRVG--KMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE---KFEKVPGKYDM  190
            ||:||       ||  |:|:|::||||  :::|| |||||:||   :|  |
Db  541  RKWGRVGSEKIGGQKLEEMS--KTEAIKEFKRLFLEKTGNSWEAWECKTNFRKQPGRFYP  598

Qy  191  LQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEMMEMKYNTK  250
            |:  :        :| : ||       :|| :|:|||  :||| :|||  ||
Db  599  LDVDYGVKKAPKRKDISE--MKSSLAPQ------LLELMKMLFNVETYRAAMEFEINMS  650

Qy  251  KAPLGKLTVAQIKAGYQSLKKIEDCIR-AGQHGRALME-----ACNEFYTRIP--HDFGL  302
           ||| |||||: :||| |:|::|::   |: |  :|          ||:|| :  | |
Db  651  EMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAASNRFFTLIPSIH----  706

Qy  303  RTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHE  362
           |||||||||||||||||||||||||||||  |||||  :|| |||||:|| ||| | ::|
Db  707  --PHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDE-SLDDKYMKLHCDITPLAHD  763

Qy  363  SYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE---KEAFREDLHNRMLLWHGSRM  419
           : ||||||| ||||||||||||||||| |:   :|||    | :||||||| ||||| ||
Db  764  SEDYKLIEQYLLNTHAPTHKDWSLELEEVFSLDRDGELNKYSRYKNNLHNKMLLWHGSRL  823

Qy  420  SNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVAL  479
           ||| ||||||||||||||| ||||||||||||||| ||||| :||:||  :|||:||||
Db  824  TNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVSKSAQYCYVDRNNPVGLMLLSEVAL  883

Qy  480  GQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGS-TVPLGPASDTGILNPDG  538
            || ||||| :||||||||||||||||| || ||||:|||||  |||| | ::|||||||
Db  884  GDMYELKKAT-SMDKPPRGKHSTKGLGKTVPLESEFVKNRDDVVVPCGKPVPSSIRSSE-  941

Qy  539  YTLNYEYIYVNPNQVRMRYLLKVQFN  565
            :|||||||||: :|||||||||:|
Db  942  --LMYNEYIYVNTSQVKMQFLLKVRFH  966
```

RESULT 14
T18600
hypothetical protein AC8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18600
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z18996
A;Accession: T18600
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-727 <WIL>
A;Cross-references: EMBL:Z83097; PIDN:CAB05448.1; GSPDB:GN00023; CESP:AC8.1
A;Experimental source: clone AC8
C;Genetics:
A;Gene: CESP:AC8.1
A;Map position: 5
A;Introns: 31/3; 271/3; 312/3; 348/1; 466/3; 545/3; 617/3; 702/2

```
Query Match       22.9%;
Best Local Similarity   31.1%;   Pred. No. 1.2e-35;
Matches  178;   Conservative  110;   Mismatches  220;   Indels  65;   Gaps  16;

Qy   19  ESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGK  78
            |:|| |          ||:||::|:|   : :
Db  189  EALAAKGGSTEPATPASASPT-----PPEAETPVLSAEGSPE-SSNKRPASAEIIEIDGE  242
```

```
Qy   79  APVDPECTAK---------------VGKAHVYCEGNDVDYDVMLNQTNLQFNNNKYLIQLLED  126
            :             |                 :   : :
Db  243  GNPDENDFAKKRRMKKEARLMEVQKKRMKKQSDLLWEYRQIFERMPYTDN----ISILRE  298

Qy  127  DAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPG  186
           |                 ||  |:  |::        ||    :|:||  |:  |
Db  299  NEQD-----------------IPEGHDRTAQDVFH---EKTKNDWIYKKHFRKRMPG  334

Qy  187  KYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ-LDLRVQELIKLICNVQAMEMMME  244
            ||  | |  |               |||    |    |: |||:| |: |: |
Db  335  MFSYVETDYSEFVGTNNGHKKKITPGSKITPGSKTLLPKSVKEVVMSIFDVENMKSALKS  394

Qy  245  MKYNTKKAPLGKLTVAQIKAGYQSLKKIED-CIRAGQHGRALMEACNEFYTRIPHDFGLR  303
            :: ::| ||||||:|| :|   |:|  : :|:                 |  |:|
Db  395  FEIDVNKMPLGRLSHQINLAFEVLNDISDLLVKLPIDASKILDFSNKFYIIIPHNFGMR  454

Qy  304  TPPLIRTQKELSEKIQLLEALGDIEIAIKLV------KTELQSPEHPLDQHYRNLHCAL  356
           |||||||||||||||||||||||||||      | :|| :|| |||||:||| ||
Db  455  VPEPIDSFHKIKEKNNMLNALLDIKFAYDQISGGDVPASTSLSI-DPVDINYRKLKCIM  512

Qy  357  RPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLHNRMLLWHG  416
           ||||| |||||||||||||||||||||||||||||||||| | |||||||||||||||||
Db  513  EPLQQGCDDWNMIHQYLKNTHGATH-DLKVELIDILKVNRDNESSKFKRHIGNRRLLWHG  571

Qy  417  SRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSE  476
           || ::| |||| |||||||| |||||||| |:||||  ::|    : |  ||
Db  572  SGKMNFAGILGQGLRIAPPEAPVSGYMFGKGVYFADMFSKSFFYCRAN-AKEEAYLLLCD  630

Qy  477  VALGQCNELLEA-NPKAEGLLQGKHSTKGLGKMAPSS-AHFVTLNGSTVPLGPA--SDTG  532
           ||| |  ||:    | |   |||||||||||| |||   ||||||  |||||    :|:
Db  631  VALGNVQQLMASKNVSRQTLPAGFQSVQGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQG  690

Qy  533  ILNPDGYTLNYEYIYVNPNQVRMRYLLKVQFN  565
            ::  |  ||||||||| :|||||||||:|
Db  691  KQNVD-YHLLYNEFIYVDVDQIQLKYLVRVKMH  722
```

RESULT 15
T08713
NAD+ ADP-ribosyltransferase homolog DKFZp566G0224.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Nov-2000
C;Accession: T08713
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: 216472
A;Accession: T08713
A;Molecule type: mRNA
A;Residues: 1-459 <ANS>
A;Cross-references: EMBL:AL050034
A;Experimental source: fetal kidney; clone DKFZp566G0224
C;Genetics:
A;Note: DKFZp566G0224.1
C;Superfamily: NAD+ ADP-ribosyltransferase

```
Query Match       22.5%;           Score 674.5;   DB 2;   Length 459;
Best Local Similarity   35.1%;   Pred. No. 3.3e-35;
Matches  170;   Conservative  91;   Mismatches  163;   Indels  61;   Gaps  15;

Qy  111  LQFNNNKYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDK  170
           :|:||||:|||||: ::|| :  ||||||||: || :|  :|:|:  :|: |:|||| :
Db    1  IENNNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDAKKDFEKKFREK  58

Qy  171  TKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ------LDLR  224
           |||||  |: | : |||| : :                            :   :|
Db   59  TKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPA  114

Qy  225  VQELIKLICNVQAMEMMEMKYNTFKKAPLGKLTVAQIKAGYQSLKKIEDCIRA-GQHGR  283
           .|:||  :  :: |  ||:  ::   ||||||: ::||:||  |  :  |   ::
Db  115  TQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSLSKQQIARGFEALEALEAALKGPTDGGQ  174

Qy  284  ALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEI-AIKLVKTELQSP  341
```

```
Db   175  SLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTV 234

Qy   342  E---HPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDG 398
Db   235  EEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHRCPTLQHIWKVNQEG 291

Qy   399  EKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSK 456
Db   292  EEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGEGIYFASENSK 347

Qy   457  SANYCFASR--LKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAH 514
Db   348  SAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL----------KSPPPGFD 393

Qy   515  FVTLNGSTVPLGPASDT--------GILNPDGY--------TLNYNEYIVYNPNQVRMR 557
Db   394  SVIARGHTEP-DPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLR 452

Qy   558  YLLKV 562
Db   453  YLLEV 457


Search completed: August 29, 2002, 07:59:24
Job time: 271 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:59:29 ; Search time 69.02 Seconds
                                                        (without alignments)
                                                        751.786 Million cell updates/sec

Post-processing: Minimum Match 08
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_71:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Title: US-09-701-586B-6
Perfect score: 2854
Sequence: 1 MSLLFLAMAPKPKPWVQTEG..........EYLIYQESQCRLRYLLEVHL 540

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	2419	84.8	459	2	T08713	NAD+ ADP-ribosyltr
2	694	24.3	635	2	T01311	NAD+ ADP-ribosyltr
3	681	23.9	1016	1	JS0428	NAD+ ADP-ribosyltr
4	674.5	23.6	996	1	S42208	NAD+ ADP-ribosyltr
5	670.5	23.5	653	2	T03656	probable NAD+ Adp-
6	670	23.5	1014	1	A29725	NAD+ ADP-ribosyltr
7	669.5	23.5	1011	1	JH0581	NAD+ ADP-ribosyltr
8	657	23.0	500	2	S26057	NAD+ ADP-ribosyltr
9	653	22.9	998	2	S31735	NAD+ ADP-ribosyltr
10	653	22.9	1013	1	S04200	NAD+ ADP-ribosyltr
11	638	22.4	994	1	A47474	NAD+ ADP-ribosyltr
12	598.5	21.0	983	1	T51353	NAD+ ADP-ribosyltr
13	598.5	21.0	1009	2	C84719	probable poly (ADP
14	578	20.3	727	2	T03657	NAD+ ADP-ribosyltr
15	534	18.7	727	2	T18600	hypothetical prote
16	393.5	13.8	538	2	T20414	hypothetical prote
17	190.5	6.7	135	2	PN0494	hypothetical prote
18	164.5	5.8	2004	2	D88948	protein ZK1005.1
19	134	4.7	1156	2	B70356	chromosome assembl
20	126	4.4	1222	2	C88504	protein B0361.3 [i
21	120.5	4.2	2044	2	AB1180	probable peptidogl
22	120	4.2	1938	1	A40997	myosin heavy chain
23	116.5	4.1	1435	2	S69632	regulatory protein
24	113.5	4.0	181	2	T03058	NAD+ ADP-ribosyltr
25	113.5	4.0	1164	2	A60234	IgA Fc receptor pr
26	113.5	4.0	1164	2	FCSOAG	IgA Fc receptor pr
27	113.5	4.0	4540	2	T30838	cytoplasmic dynein
28	113	4.0	2078	2	T25400	hypothetical prote
29	110	3.9	1534	2	A56734	ribosome receptor,
30	110	3.9	2288	2	T29999	hypothetical prote
31	109.5	3.8	446	2	A46335	gag polyprotein -
32	108.5	3.8	388	1	QOTV	transforming prote
33	108.5	3.8	1233	2	T30534	chromosome segrega
34	108.5	3.8	1871	2	S27938	chromosome segrega
35	108	3.8	359	2	T12540	hypothetical prote
36	108	3.8	1092	2	T13717	carbamoyl-phosphat
37	108	3.8	3259	2	S52300	giantin - human
38	108	3.8	3225	2	A55539	giantin - human
39	107.5	3.7	716	2	G69708	hypothetical prote
40	107	3.7	1186	2	T76998	chromosome segrega
41	106.5	3.7	568	2	T34522	hypothetical prote
42	106	3.7	725	1	JC5016	hyaluronan recepto
43	106	3.7	932	2	S62555	protoplast regener
44	106	3.7-	1199	2	T29145	hypothetical prote
45	106	3.7	2176	2	T13806	toucan gene protei

ALIGNMENTS

RESULT   1
T08713
NAD+ ADP-ribosyltransferase homolog DKFZp566G0224.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Nov-2000
C;Accession: T08713
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16472
C;Accession: T08713
A;Molecule type: mRNA
A;Residues: 1-459 <ANS>
A;Cross-references: EMBL:AL050034
A;Experimental source: fetal kidney; clone DKFZp566G0224
C;Genetics:
A;Note: DKFZp566G0224.1
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match   84.8%; Score 2419; DB 2; Length 459;
Best Local Similarity   99.8%; Pred. No. 1.4e-161;
Matches 458; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

```
Qy     82 IENNNNKFYIIQLLQQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREKTK 141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 IENNNNKFYIIQLLQQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREKTK  60

Qy    142 NNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRYQPCSLDPATQRKLIT 201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 NNWAERDHFVSHPGKYTLIEVQAEDEAQEAAVVKVDRGPVRTVTKRVQPCSLDPATQKLIT 120

Qy    202 NIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELS 261
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 NIFSKEMFKNTMALMDLDVKRMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELS 180

Qy    262 SHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQOALQAVSEQEKTVEEVPHP 321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 SHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQOALQAVSEQEKTVEEVPHP 240

Qy    322 LDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVNQEGEEDRFQAHS 381
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 LDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVNQEGEEDRFQAHS 300

Qy    382 KLGNRKLLIWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGYVIGMKCGAH 441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 KLGNRKLLIWHGTNMAVVAAILTSGLRIMPHSGGRVGEGIYFASENSKSAGYVIGMKCGAH 360

Qy    442 HVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVV 501
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 HVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVV 420

Qy    502 PQGQPVPCPEFSSSTFSQSEVYLIYQESQCRLRYLLEVHL 540
          ||||||||||||||||||||||||||||||||||||||||
```

Db 421 PQCQOPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 459
         ||||||||||||||||||||||||||||||||||||||||

RESULT 2

T01311
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana
N:Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19
A:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C:Accession: T01311
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-635 <KAL>
R:Kalicki, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana T14P8.
A:Reference number: Z14290
R:Lepiniec, L.; Babiychuk, E.; Kushnir, S.; Van Montagu, M.; Inze, D.
FEBS Lett. 364, 103-108, 1995
A:Title: Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(ADP-r
A:Reference number: S65662
A:Accession: S65662; MUID:95269779
A:Molecule type: mRNA
A:Residues: 1-115; 'GT',116-635 <LEP>
C:Cross-references: EMBL:Z48843; NID:g853721; PIDN:CAA88288.1; PID:g853722
C:Genetics:
A:Map position: 4
A:Introns: 28/3; 116/1; 162/3; 190/1; 268/3; 312/1; 330/3; 355/2; 371/1; 387/3; 408/3; 4
A:Gene: PARP
A:Note: T14P8.19
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger

Query Match   24.3%;   Score 694;   DB 2;   Length 635;
Best Local Similarity 34.2%;   Pred. No. 1.2e-40;
Matches 185; Conservative 97; Mismatches 221; Indels 38; Gaps 18;

Qy   18 TEGEPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRV-DPTCPLSSNPGTQVYED---- 72
             |||||:||||||| : ||::||:::|: ||::::|||||: : ||:||:: :::
Db 107 SNAPVKSSNDEAEDDNGFEEEKKEEKIVTAIKKGAAVLDQWIPDEIKSQYHVLQRGDDV 166

Qy   73 YNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK 130
         |:|||||||:|:|||||||::| ||:||||||||||||:||||::| :|::|:|:||
Db 167 YDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVYTRWGRVGKGQSKLDGYPDSWDRAIE 226

Qy 131 DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQP- 189
         ::::|:::|||| | :|:::|:|||:||:|:|||::|:::| : |: :||:| :::|
Db 227 IFTNKFNDKTKNYWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNND---IPSSSSEVKPE 283

Qy 190 -CSLDPATQKLLTNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEEALK 248
          ||||: :::|::::: |:::| ||::||:|||||||||:|||:|:|:|||::|:
Db 284 QSKLDTRVAKFISLICNVSMMAQHMMEIGYNANKLPLGKISKSTISKGYEVLKRISEV1- 342

Qy 249 GPTDGGQSLEELSSHFYTVIPHNFGHSQPPP- INSPELLQAKKDMLLVLADIELAQALQ 306
         ::::::::||||||:|:||||||:||:::| ||::|||:||:|::||: ||:|::|:|:
Db 343 -DRYDRTRLEELSEGEFVTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLL 401

Qy 307 AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQ1-GSNHRCPTLQ-- 363
         :| :::: : :||||: :|:|:::::::|:|:::: |:::::::| :::: : |:|
Db 402 SVDPGLQ-----DDPLYHYQQLNCGLTPVGNDSEEFSMVANYMENHHAKTHSGYTVEIA 456

Qy 364 HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKG 419
         :|:: ::: ::|||| :::|:|:|||||::: | :| ::|||:| || : ||:||||
Db 457 QLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKG 516

Qy 420 IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINYDNPSLKSPPPGFDSVIAR 479
         :|||: ||| | |:|:| | : ::|:|:||:||||::::::| |: | |||::: :|
Db 517 VTFADMFSKSANYCYA-NTGAND-GVILLCEVALGDMNELLYSDYNADNLPGKLSTGGV 574

Qy 480 GHTEPDTQDTELELDGDQQVVVPQGQPV - PCPEFSSSTFSQSEYLIYQESQCRLRYLLE 537
         |:| : ::||:|:|: :|::|| :||| |||||||||||||:||:|||||||||||:|
Db 575 GKTAPNPSEAQTLE-DG--VVVPLGKPVERSC---SKGMLIYNEYIVYNVEQIKMRYVIQ 628

Qy 538 V 538
         |
Db 629 V 629

RESULT 3

JS0428
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine
N:Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate
A:Species: Bos primigenius taurus (cattle)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C:Accession: JS0428; S00328; A30458
A:Status: translated from JIPID; February 1990
R:Saito, I.
submitted to JIPID, February 1990
A:Reference number: JS0428
A:Accession: JS0428
A:Molecule type: mRNA
A:Residues: 1-1016 <SAI>
A:Experimental source: thymus
R:Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; T
Eur. J. Biochem. 171, 571-575, 1988
A:Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int
A:Reference number: S00328; MUID:88151954
A:Molecule type: mRNA
A:Residues: 648-714; 838-904 <TAN>
A:Accession: S00328
A:Cross-references: EMBL:X06986
A:Accession: A30458
A:Molecule type: protein
A:Residues: 658-685; 689-696; 893-901 <TA2>
C:Superfamily: NAD+ ADP-ribosyltransferase
C:Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-loo
F:21-51/Region: zinc finger
F:128-165/Region: zinc finger
F:200-220/Region: helix-turn-helix motif
F:224-231/Region: nuclear location signal
F:250-270/Region: helix-turn-helix motif
F:494-501/Region: nucleotide-binding motif A (P-loop)
F:890-903/Region: nucleotide binding #status predicted

Query Match   23.9%;   Score 681;   DB 1;   Length 1016;
Best Local Similarity 33.9%;   Pred. No. 1.8e-39;
Matches 188; Conservative 99; Mismatches 205; Indels 62; Gaps 24;

Qy   7 AMAPKPK----PWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLS 62
         : | ::: ||::||:|||||:|||||||:|||:::|: :
Db 494 AVGPKGKSGAAPSKKSKGPVKEEG--TNKSEKRMKLTLKGGAA----VDPDSGLE 542

Qy   63 SNPGTQVYED----YNCTLNQTNIENNNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV-GQ 116
         ||: :::: :||||||:|:::| :::||::|||:||:||||:|:
Db 543 HN--AHVLEKGGKVFSATLGIVDIVYKGINSYYKLQLLEDDKESRYWIFRSWGRVGTVIGS 600

Qy 117 SKINHFTRLEDAKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKV 175
         |:::|: :|: |:|::::::||:|| | |:|:::|:|||:|:|:| |::|:::| : |
Db 601 NKLEQMPSKEDAIEIHPMKLYEEKTGNAWHSK-NETKHPRKFYPLEIDYGQDE--EAVKKL 657

Qy 176 DRGPVRTVTKRVQPCSLDPATQKLLTNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIAR 235
         : ||:| :::| ||||: :::|::::: |:::: |::||:|||||||||:|||:|
Db 658 TVNP-GTKSKLPKP-----VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQA 710

Qy 236 GFEALEALEEAL-KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL 294
         ::|:|: ||:|||::::::||||||||:|:|||||||:||:::||:|||||:||:|::||
Db 711 AYSILSEVQOALSQGSSD--SHILDLSNRFYTLLIPHDFGMKKPPLLNNANSVQAKVEMLD 768

Qy 295 VLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQ- 353
         :|:||| :::::| :: :::: : :||||: :|:|:::::::|:|:::: |:::::
Db 769 NLLDIEVAYSLLRGGSDSSKD----PIDVNYEKLKTDIKVVDKDSEPAEIIRKVKNTH 824

```
Qy   354  GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH  411
Db   825  ATTHNAYDLEVVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPP  884

Qy   412  -----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGTMFLGEVALGREHHINTDNP  464
Db   885  EAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGILLLGEAALGNMYELKHAR-  938

Qy   465  SLKSPPPGFDSVIARGHTEEDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLI  524
Db   939  HISKLPKGHSVKGLGKTTPDPS--ASITVDG--VEVPLGTGI-SSGVNDKCLLNEYIV   993

Qy   525  YQESQCRLRYLLEV  538
Db   994  YDIAQVHLKYLLKL  1007
```

RESULT   4

S42208

NAD+ ADP-ribosyltransferase

N:Alternate names: poly(ADP-ribose) polymerase
C:Species: Sarcophaga peregrina
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: S42208; S71496
R:Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurat
Eur. J. Biochem. 220, 607-614, 1994
A:Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from Sarc
A:Reference number: S42208; MUID:94170813
A:Accession: S42208
A:Molecule type: mRNA
A:Residues: 1-996 <MAS>
A:Cross-references: EMBL:D16482; NID:g473742; PIDN:BAA03943.1; PID:g538248
A:Accession: S71496
A:Molecule type: protein
A:Residues: 170-188;721-736;813-819;879-885 <MAX>
A:Superfamily: NAD+ ADP-ribosyltransferase
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger
F:1-369/Domain: DNA binding #status predicted <DNA>
F:370-507/Domain: auto-modification #status predicted <AMO>
F:508-996/Domain: NAD binding #status predicted <NAD>

Query Match   23.6%;   Score 674.5;   DB 1;   Length 996;
Best Local Similarity 33.6%;   Pred. No. 5.1e-39;
Matches 179;   Conservative 104;   Mismatches 195;   Indels 55;   Gaps 22;

```
Qy    32  EEDPFRSTAEALKAIPAEKRI-----IRVDPTCPLSSNPGTQVY------EDYNCTLNQTN   81
Db   490  ESKSKSKSKSIYTKSVPKSMTLKIKDGIAVDPDSGLED--VAHVYVSRNKEKYNVVLGITD  547

Qy    82  IENNNNKFYIIQLLQD--SNRFFTCWNRWRGVG-EVGQSKINHFTRLEDAKKDFEKKFRE   138
Db   548  IQKNKNSFYKLQLLESDMKNRFWV-FRSWGRIGTTIGGNKLDNFSNLVDAIVQFKELYLE   606

Qy   139  KTKNNNAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQ   197
Db   607  KSGNHFENRENFVKVAGRMYPIDIYAEDS-----KIDLSAEHDIKSKL-PLSV----    655

Qy   198  KLITNIFSKEMFKNTMALMDLLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSL  257
Db   656  DIIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLTEIYELIQ---GGTNA   712

Qy   258  E--ELSSHFYTVIPIHNEFRGHSQOPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTV  315
Db   713  KFIDATNRFYTLIPIHNEFRGHTQOSPPLLDTTEQVEOLRQMLDSLIEIECAYSLLQTEDSKADI  772

Qy   316  EEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEG   372
Db   773  ----NPIDKHYEQLKTKLEPLDKNSEEYILLQKYVKNTHAETHKLIYDLEVVDIFKVAROG  828

Qy   373  EEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH-----SGGRVGKGIYFASENSK  428
```

```
Db   829  EARRYKPFKLHNRRLLWHGSRTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADMVSK   888
Qy   429  SAGYVIGMKCGAHH--VGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEDP   486
Db   889  SANYC----CTSHHNSTGLMLLSEVALGDMMECTAAKVTKLPNDK-HSCFGRGRTMPNP   943
Qy   487  TQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVH   539
Db   944  SESIIRE-DG--VEIPLGKPITNDSLKSSLL-YNEFIIYDIAQVNIQYMLRMN   992
```

RESULT   5

T03656

probable NAD+ ADP-ribosyltransferase (EC 2.4.2.30) -- maize

C:Species: Zea mays (maize)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Oct-1999
R:Babylchuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Mo
submitted to the EMBL Data Library, November 1997
C:Description: Higher plants possess two poly(ADP-ribose) polymerases.
A:Reference number: Z14991
A:Accession: T03656
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-653 <BAB>
A:Cross-references: EMBL:AJ222588; NID:e1264090; PIDN:CAA10888.1; PID:e1264091
A:Genetics:
A:Gene: PARP
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match   23.5%;   Score 670.5;   DB 2;   Length 653;
Best Local Similarity 35.4%;   Pred. No. 5.3e-39;
Matches 170;   Conservative 93;   Mismatches 186;   Indels 31;   Gaps 14;

```
Qy    71  EDYNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWRGRVGEVGQSKINHFTRLED-A  128
Db   188  EIYDATLNQTNVGDNNNKFYIIQVLESDAGGSPMVVNRWGRVGVRGQDKLHGPSPTRDOA   247

Qy   129  KKDFEKKFEREKTRKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ  188
Db   248  IYEFEGKIPHNKTNNHWSDRKNFRCYAKKYTWLEMDYGETEKE---IEKG--SITDQIK   300

Qy   189  PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK   248
Db   301  ETKLETRIAQFISLICNISMMKQRMVEIGYNAEKLPLGKLRATILKGYHVLKRISDVIS   360

Qy   249  GPTDGGQSLEELSSHFYTVIPIHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ  306
Db   361  --KADRRHLEQLTGEFYTVIPIHDFGFRKMREFFIDTPQKLKAKLEMVEALGEIEIATKL-  417

Qy   307  AVSEQEKTVEEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ-  363
Db   418  ----LEDDSSDQDDPLYARYKQLHCDFFPLEADSDEYSMIKSLRNTHGKTHSGYTVDIV   473

Qy   364  HIWKVNQEGEEDRFQAHSKLLWHGTNMAVVAAILTSGLRIMPH--SGGRVGKG   419
Db   474  QIFKVSRHGETERFQKFASTRNRMLLWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGKG   533

Qy   420  IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR   479
Db   534  VIFADMFSKSANYCIASE--ACKSGVLLCEVALGDMNELLNADYDANNLPKGKLRSKGV   591

Qy   480  GHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVH   539
Db   592  GQTAPNMV-ESKVADDG--VVVPLGEPKQEPS-KRGGLLYNEIVIVYNVDQIRMRYVLHVN  647
```

RESULT   6

A29725

NAD+ ADP-ribosyltransferase (EC 2.4.2.30), nuclear - human

N:Alternate names: poly (ADP-ribose) polymerase; poly (ADP-ribose) synthetase; poly(A
C:Species: Homo sapiens (man)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29725; A28498; A39976; A26901; I38096; B33321; A33321; A35635; A61559; S14
R;Uchida, K.; Morita, T.; Sato, T.; Ogura, T.; Yamashita, R.; Noguchi, S.; Suzuki, H.; N
R;Kurosaki, T.; Ushiro, H.; Mitsuuchi, Y.; Suzuki, S.; Matsuda, M.; Katunum
J. Biol. Chem. 262, 15990-15997, 1987
A;Title: Primary structure of human poly(ADP-ribose) synthetase as deduced from cDNA se
A;Reference number: A28498; MUID:88058958
A;Accession: A28498
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-16, 'E', 18-211, 'K', 213-236, 'R', 238-366, 'H', 369-1014 <KUR>
A;Cross-references: GB:J03473
R;Cherney, B.W.; McBride, O.W.; Chen, D.; Alkhatib, H.; Bhatia, K.; Hensley, P.; Smulson
Proc. Natl. Acad. Sci. U.S.A. 84, 8370-8374, 1987
A;Title: cDNA sequence, protein structure, and chromosomal location of the human gene fo
A;Reference number: A39976; MUID:88068596
A;Accession: A39976
A;Molecule type: mRNA
A;Residues: 1-49, 'D', 51-612, 'Q', 614-907, 'Y', 909-939, 'R', 941-979, 'I', 981-1014 <CHE>
A;Cross-references: GB:J03030
R;Suzuki, H.; Uchida, K.; Shima, H.; Sato, T.; Okamoto, T.; Kimura, T.; Miwa, M.
Biochem. Biophys. Res. Commun. 146, 403-409, 1987
A;Title: Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and expression
A;Reference number: A26901; MUID:87298455
A;Accession: A26901
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 41-610, 'N', 612-880;921-1014 <SU2>
A;Residues: 1-40 <RES>
R;Ogura, T.; Nyunoya, H.; Takahashi-Masutani, M.; Miwa, M.; Sugimura, T.; Esumi, H.
Biochem. Biophys. Res. Commun. 167, 701-710, 1990
A;Title: Characterization of a putative promoter region of the human poly(ADP-ribose) po
A;Reference number: I38096; MUID:90211250
A;Accession: I38096
A;Molecule type: DNA
A;Cross-references: EMBL:X16674; NID:g510112; PIDN:CAA34663.1; PID:g1017423
R;Auer, B.; Nagl, U.; Herzog, H.; Schneider, R.; Schweiger, M.
DNA 8, 575-580, 1989
A;Title: Human nuclear NAD(+) ADP-ribosyltransferase(polymerizing): organization of the
A;Reference number: A33321; MUID:90091744
A;Accession: B33321
A;Molecule type: DNA
A;Residues: 38-43;93-98;132-137;204-209;237-242;276-281;335-340;384-389;431-436;512-517;
A;Cross-references: GB:M29544; GB:M22953
A;Note: the authors translated the codon GTG for residue 54 as Glu
A;Note: these fragments represent intron-exon boundaries
A;Accession: A33321
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 16-66;96;121-159,'D',161-167 <AU2>
A;Note: these fragments represent a zinc finger-containing DNA-binding region
A;Cross-references: GB:M29544; GB:M22953
A;Note: the authors translated the codon GTG for residue 54 as Glu
A;Note: these fragments represent intron-exon boundaries
A;Accession: A33321
R;Gradwohl, G.; Menissier de Murcia, J.; Molinete, M.; Simonin, F.; Koken, M.; Hoeijmake
Proc. Natl. Acad. Sci. U.S.A. 87, 2990-2994, 1990
A;Title: The second zinc-finger domain of poly(ADP-ribose) polymerase determines specifi
A;Reference number: A35635; MUID:90221155
A;Accession: A35635
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 12-26,'T',28-66;116-166 <GRA>
R;Schneider, R.; Auer, B.; Kuehne, C.; Herzog, H.; Klocker, H.; Burtscher, H.J.; Hirsch-
Eur. J. Cell Biol. 44, 302-307, 1987
A;Title: Isolation of a cDNA clone for human NAD (+): protein ADP-ribosyltransferase.
A;Reference number: A61559; MUID:88082900
A;Accession: A61559
A;Molecule type: mRNA

A;Residues: 381-420;682-710 <SCH>
R;Yokoyama, Y.; Kawamoto, T.; Mitsuuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Tera
Eur. J. Biochem. 194, 521-526, 1990
A;Cross-references: EMBL:X56140; NID:g935286; PIDN:CAA39606.1; PID:g825702
A;Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.
A;Reference number: S14010; MUID:91099327
A;Accession: S14010
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-95 <YOK>
C;Cross-references: EMBL:X56140; NID:g935286; PIDN:CAA39606.1; PID:g825702
C;Comment: This protein can ADP-ribosylate itself as well as other proteins.
C;Genetics:
A;Gene: GDB;ADPRT; PPOL
A;Map position: 1q41-1q42
C;Superfamily: NAD+ ADP-ribosyltransferase; GDB:1119508; OMIM:173870
C;Cross-references: GDB:1119508; OMIM:173870
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltrans

Query Match   23.5%;   Score 670; DB 1; Length 1014;
Best Local Similarity 33.6%;   Pred. No. 1.1e-38;
Matches 186; Conservative 101; Mismatches 195; Indels 72; Gaps 23;

```
Qy    7 AMAPKPKPWVQTEGPEKKKGR------QAGREEDP---FRSTAEALKAIPAEKRIIRVDPI  58
          |::| ||  |::|                    |     :| :    |: |     |:
Db  502 ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLE-------------- 547

Qy   59 CPLSSNPGTQVYEDYCTLNQTNIENNNNKFYIIQLLQD--SNRFTTCWNRWGRVGEV-G  115
           |::|  ||  |:|| ||:    ||   ||   ::|    : ||||||||||||| |
Db  548 ------KGGKV---FSATLGLVDIVKGTNSYYKLQLLEDDKENRYW-IFRSWGRVGTVIG  597

Qy  116 QSKINHFTRLEDAAKKDFEKKFEREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVK  174
          ||:  || |||  |  || |  | |  |   |:  || ::||| |||      |  | |
Db  174 AAYSILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQOAKVEMLD  766

Qy  175 VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA  234
          :::  |  | |:|  | | |  |: |    | :|  | : ||   |: ::| ||  ::
Db  234 NLLDIEVAYSLLRGGSDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEIIRKVVKNTH  822

Qy  235 RGFEALEALEBALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL  294
          |||      |  |   |  | |: :   | |: |    ||   |     |    | : |
Db  767 NLLDIEVAYSLLRGGSDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEIIRKVVKNTH  822

Qy  295 VLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-  353
          || | |  | |::   |  | :|    || :   |  |  |  |     ::    | |
Db  354 GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH  411

Qy  354 GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH  411
Db  823 ATTHNAYDLEVTDIFKTEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPP  882

Qy  412 ----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH----VGTMFLGEVALGREHHINTDNP  464
Db  823 ATTHNAYDLEVTDIFKTEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPP  882

Qy  465 SLKSPPPGFDSVIARGHTEPDPTODFELEDLGQOVVWPQGQPVPCPEFSSSTFSQSEYLI  524
Db  883 EAPVTGYMFGKGIYFADMVSKSANY----CHTSQGDPIGILLLGEVALGNMYELKHAS-  936

Qy  525 YQESQCRLRYLLEV 538
Db  937 HISKLPKGHSVKGLGKTTPDPS--ANISLDG--VDVPLGTGI-SSGVNDTSLLYNEYIV  991

Db  992 YDIAQVNLKYLLKL 1005
```

RESULT 7
JH0581
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - chicken
N;Alternate names: poly(ADP-ribose) synthase
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JH0581
R;Ittel, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.
Gene 102, 157-164, 1991

```
A;Title: Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence and c
A;Reference number: JH0581
A;Accession: JH0581; MUID:91340148
A;Molecule type: mRNA
A;Residues: 1-1011 <ITT>
A;Cross-references: EMBL:X52690; NID:g63742; PIDN:CAA36917.1; PID:g63743
C;Comment: This protein is a chromatin-bound enzyme.
C;Comment: This enzyme catalyzes DNA-dependent post-translational modifications of vario
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc fi

Query Match    23.5%;   Score 669.5;  DB 1;  Length 1011;
Best Local Similarity  33.2%;  Pred. No. 1.2e-38;
Matches 186;  Conservative 100;  Mismatches 205;  Indels 69;  Gaps 24;

Qy   14 PW---VCTEGPE------------------KKKGRQAGREEDPFPRSTAEALKAIPAEKRIIRVDPT 58
                                                                        ||
Db  477 PWGAEVKTEHQEVAVDGKCCSKRPANMKSAGKVKEEQGPSKSEKRMKLTV---KGGAAVDPD 533

Qy   59 CPLSSNPGTQVYED---------YNCTLNQTNIENNNNKFYIIQLLQD--SNRFFTCWNRWGRVG 112
                                                                        |||
Db  534 SGLEDS--AHVFEKGGKIFSATLGLVDIVKGTNSYIKLQLLEDDRESRYWV-FRSWGRVG 590

Qy  113 EV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQE 170
Db  591 TVIGSNKLEQMPSKEDAVEHFLNLYEEKTGNSWHSK-NFTKYPKFYPLEIDYGQDEAV 649

Qy  171 AVVKVDRGPVRVTVKRVQPCSLDPATQKLIITNIFSKEMFKNTMALMDLDVKKMPLGKLSK 230
Db  650 RKLTVSAG---TKSKLAKP------IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSK 700

Qy  231 QQIARGFEALEALEEALKGPTDGGQSLE--ELSSHFYTVIPHNFGHSQPPPINSPELLQA 288
Db  701 RQIQSAYSILNEVQQAV---SDGGSESQOIDLSNRFYTLIPHDFGMKKPPLLSNLEYIQA 757

Qy  289 KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQT 348
Db  758 KVQMLDNLLDIEVAYSLLRGGNEDGDKD----PIDINYEKLRIDIKVVDKDSEEAKIIKQ 813

Qy  349 YLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSG 405
Db  814 YVKNTHAATHNAYDLKYVEIFRIEREGESQRYKPFKQLHNRQLLWHGSRTTNPAGILSQG 873

Qy  406 LRIMPH--------SGGRVGKGIYFASENSKSAGYVIGMKC--GAHVGYMFLGEVALGREHH 458
Db  874 LRIAPPEAPVIGYMFGKGIYFRDMVSKSANY-----CHTSQADPIGLILLGEVALGNMYE 928

Qy  459 INTDNPSLKSPPPGFDSVIARGHTEDPDTQDTELELDGQQVVVPQGQVPCPEEFSSSTFS 518
Db  929 LKNASHITKL-PKGKHSVKGLGKTAPDTATT--TLDG--VEVPLGNGI-STGINDTCLL 982

Qy  519 QSEYLIYQESQCRLRYLLEV 538
Db  983 YNEYIVYDVAQVNLKYLLKL 1002
```

RESULT 8
S26057
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 18-Jun-1999
C;Accession: S26057; S78453; 152331
R;Thibodeau, J.; Gradwohl, G.; Dumas, C.; Clairoux-Moreau, S.; Brunet, G.; Penning, C.;
Biochem. Cell Biol. 67, 653-660, 1989
A;Title: Cloning of rodent cDNA encoding the poly(ADP-ribose) polymerase catalytic domai
A;Reference number: 152331; MUID:90027702
A;Accession: S26057
A;Molecule type: mRNA
A;Residues: 1-500 <THI>
A;Cross-references: EMBL:X65497; NID:g56849; PIDN:CAA46478.1; PID:g56850

```
A;Reference number: S78453
A;Accession: S78453
A;Molecule type: mRNA
A;Residues: 1-124, 'H', 126-127, 'A', 129-238, 'D', 240-500 <POT>
A;Cross-references: EMBL:X65497
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltrans

Query Match    23.0%;   Score 657;  DB 2;  Length 500;
Best Local Similarity  34.4%;  Pred. No. 3.2e-38;
Matches 177;  Conservative 96;  Mismatches 190;  Indels 52;  Gaps 21;

Qy   48 AEKRI-------IRVDPTCPLSSNPGTQVYED----AHVLEKGCKVFSATLGLVDIVKGTNSYIKLQLLE 96
                ||||                              ||:|::
Db    5 SEKRMKLITLKGGAAVDPDSGLEHS--AHVLEKGCKVFSATLGLVDIVKGTNSYIKLQLLE 62

Qy   97 -DSNRFFTCWNRWGRVGEV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHP 154
Db   63 SDKESRYWIFRSWGRVGTVIGSNKLEQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTKYP 121

Qy  155 GKYTLIEVQ-AEDEAQEAVVKVDRGPVRVTVKRVQPCSLDPATQKLIITNIFSKEMFKNTM 213
Db  122 KKFYPLEIDYGQDE-EAVKKL---AVKPGTK----SKLPKVQELVGMIFDVESMKKAL 172

Qy  214 ALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFG 273
Db  173 VEYEIDLQKMPLGKLSRQIQOAYSILSEVVQQAVSQGSSESQOIL-DLSNRFYTLIPHDFG 231

Qy  274 HSQPPPINSPELLQOAKKDMLLVLADIELAQOALQAVSEQEKTVEEVPHPLDRDYQLLKCQL 333
Db  232 MKKPPLLNNTDSVQAKVEMLDNLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDI 287

Qy  334 QLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLW 390
Db  288 KVVDRDSEEAEVIRKYVKNTHAATTHNAYDLEVIDIFKIEREGESQRYKPFRQLHNRRLLW 347

Qy  391 HGTNMAVVAAILTSGLRIMPH-----SGGRVGKGIYFASENSKSAGYVIGMKCCGAHH---V 443
Db  348 HGSRTTNPAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPI 402

Qy  444 GYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEDPDTQODTELELDGQQVVVPQ 503
Db  403 GLILLGEVALGNMYELKHAS-HISKLPKGKHSVKGLGKTAPDPS--ASTLDG--VEVPL 457

Qy  504 GQPVPCPEEFSSSTFSQSEYLIYQESQCRLRYLLEV 538
Db  458 GTGIP-SGVNDTCLLYNEYIVYDIAQVNLKYLLKL 491
```

RESULT 9
S31735
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - African clawed frog (fragment)
N;Alternate names: poly ADP-ribose polymerase
C;Species: Xenopus laevis (African clawed frog)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S31735; PN0495
R;Saulier-le Drean, B.M
submitted to the EMBL Data Library, May 1992
C;Accession: S31735; PN0495
R;Ozawa, Y.; Uchida, K.; Uchida, M.; Ami, Y.; Kushida, S.; Okada, N.; Miwa, M.
Biochem. Biophys. Res. Commun. 193, 119-125, 1993
A;Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)polymera
A;Reference number: PN0494; MUID:9327538
A;Accession: PN0495
A;Molecule type: mRNA
A;Residues: 1-998 <SAU>
A;Cross-references: EMBL:Z12139; NID:g649967; PIDN:CAA78126.1; PID:g1334661

C;Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pentosyl

Query Match   22.9%;   Score 653;   DB 2;   Length 998;
Best Local Similarity   32.5%;   Pred. No. 1.6e-37;
Matches 177;   Conservative 102;   Mismatches 214;   Indels 52;   Gaps 20;

```
Qy      10 PKPKPWPVQTEGPEKKGRQAGREEDFPRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQV 69
              |  | |      | :: :|  |||  |:| :|||   ||| :|| | ||||||||||
Db     480 PSSGPVAGKSSGSVKEEKGSNKSEKRKMKLTVKGGAAIDPDSEL---EDSCHVLETGG--- 533

Qy      70 YEDYNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEV-GQSKINHFTRLED 127
              :| :::|  :  :  ::  :::||:  |  ::| | | ||  |  |||||||| |||
Db     534 -KIFSATLGLVDITRGTNSYKLQLLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKED 592

Qy     128 AKKDFEKKFREKTKNNWAAERDHFVSHPGKRYTLLIEVQAEDEAQGEAVVKVDRGPVRTVTKRV 187
              |::|  |:|| | :   | ||| |||| || |||||:||:|:|  |:|  | | ||  |:||
Db     593 AIEHFLNLYQDKTGNAW-HSPNFTKYPKFYPLEIDYGQE-EDVVKLSVG-AGTKSKLA 649

Qy     188 QPCSLDPATOKLITNIFSKEMFKNTNALMDLDVKKMPLGKYLSKQQIARGFEALEALEEAL 247
              ||: || | |:|:|:| |::  | :|||| |:|:||||||:| |||  :| :||: ||
Db     650 KP------VQELIKLIFDVESMKKAAVEFEIDLQKMPLGKLSKRQIOSAYSILSQVQQAV 703

Qy     248 KGPTDGGQSLEE------LSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELA 302
              | || | :||        | ::||| |||  |||| | | |:|::|  :|||:||||:||
Db     704 ------SESLSEARLLDLSNQFYTLIPHDFGMKKPPLLNNLEYIIQARVQMLDNLLDIEVA 757

Qy     303 QAL--QAVSEQEKTVEEVHPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRC 359
              ||   || :|::|||  ||  |||||  ||: ||  ||||:| ||| |||||:| ||||||
Db     758 YSLLRGGADDGEK------DPIDVKYEKIKIDIKVVAKDSEESRICDYVKNTHADTHNA 811

Qy     360 PTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH-----SG 413
              |||   ||||||||||||||||||||||||||||| |:|||||:||||||||      ||
Db     812 YDLEVLEIFKIDREGEYQRYKPFKQLHNROLLWHGSRTTNFAGILSQGLRIAPPEAPVTG 871

Qy     414 GRVGVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPGF 473
              ||| ||||||||||||||||||||||||||| ||||||||||| |:||:||||:|||||
Db     872 YMFGKGIYFADMVSKSANYCHAMP---GSPIGILLGEVALGNMHELKAAASQTTKL-PKGK 928

Qy     474 DSVIARGHTEPDPTQDTELELDGQQVVVPQGGQPVPCPEFSSSTFSQSEYLIYQESQCRLR 533
              |:| |||||||||||||||||||::: |||  ||||||||||||||||||| ||||||||
Db     929 HSVKGLGRTAPDPS--ATVQLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLK 983

Qy     534 YLLEV 538
              |||:|
Db     984 YLLKL 988
```

Qy   2 SLLFLAMAPRPKPW--------VQTEGPE---------------KKKGRQAGREEDFPRSTAEALKAIP 47

Query Match   22.9%;   Score 653;   DB 1;   Length 1013;
Best Local Similarity   32.3%;   Pred. No. 1.7e-37;
Matches 186;   Conservative 102;   Mismatches 211;   Indels 76;   Gaps 23;

RESULT 10
S04200
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999
C;Accession: S04200
R;Huppi, K.; Bhatia, K.; Siwarski, D.; Klinman, D.; Cherney, B.; Smulson, M.
Nucleic Acids Res. 17, 3387-3401, 1989
A;Title: Sequence and organization of the mouse poly (ADP-ribose) polymerase gene.
A;Reference number: S04200; MUID:89263780
A;Accession: S04200
A;Molecule type: DNA
A;Residues: 1-1013 <HUP>
A;Cross-references: EMBL:X14206; NID:g49893; PIDN:CAA32421.1; PID:g49894
C;Genetics:
C;Map position: 1
A;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc fi

---

RESULT 11
A47474
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: A47474
R;Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.
Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993
A;Title: Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase: leucine zip
A;Reference number: A47474; MUID:93234521
A;Accession: A47474
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-994 <UCH>
A;Note: sequence extracted from NCBI backbone (NCBIN:129703, NCBIP:129704)
A;Cross-references: GB:D13806; GB:D13807; GB:D13808; NID:g303545; PIDN:BAA02964.1; PI
A;Gene: FlyBase:Parp
A;Genetics:
A;Cross-references: FlyBase:FBgn0010247
A;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase

```
Qy      48 AERRI-------IRVDPTCPLSSNPGTQVYED----YNCTLNQTNIENNNNKFYIIQLLQ 96
              | ||        ||||||||||||||||||     |||||||| :|:::::::|||||:
Db     468 SLQDLLSAHSLSPWGAEVKAEPGEVVAPRGKSAAPSKKSKGCFKEEGVNKS-------- 518

Qy      97 DSNRF-FTCWNRWGRVGEV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAAERDHFVSHP 154
              ||||| | | ||  ||| | |||||||| ||||||||:|||||||| |||||||| ||:|
Db     519 -ERRMKLTLKGGAAVPDPDSGLEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYKLQLLE 575

Qy     155 GKRYTLLIEVQ-AEDEAQGEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTM 213
              || |||||||  |:|||  |:|  | | || |  ||| |:|| | |:|:|:|:|||:|::|
Db     576 DDKESRYWIFRSWGRLGTVIGSNKLEOMPSKEEAVEQFMKLYEEKTGNAWHSK-NFTKYP 634

Qy     214 ALMDLDVKKMPLGKYLSKQQIARGFEALEALEEALKGPTDGGQSLEEELSSHFYTVIPHFG 273
              || | |:|:||||| | ||| ||||||: ||  ||||||||  |:|:||| ||||||||
Db     635 KKFYPLEIDYGQDE--EAVVKKLTVKP-GTKKSKLPKP------VQELVGMIFDVDSMKRAL 685

Qy     274 HSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVHPHPLDRDYQLLKCQL 333
              ::| | |||:|:|| :|::|||:|||:|| |||   ||::| |||| ||:|||| ||:
Db     686 VEYEIDLQKMPLGKLSRRQIOAAYSILSEVQOPVSQGSGSESQII-DLSNRFYTLIPHDFG 744

Qy     334 QLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLW 390
              :|  ||||:|||| |||: | |||||| || |  ||||:||||||||||||||||||||
Db     745 MKKPPLLNNADSVQOAKVEMLDNLLDIEVAYSLLRGGSDDSSKD---PIDVNEKLKTDI 800

Qy     391 HGTNMAVVAAILTSGLRIMPH-----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---V 443
              |||:|||||:||||||||||      ||| |||||||||||||||||||||||||   |
Db     801 KVVDRDSEEAEVIRKYVKNTHATTHNAVDLEVIDIFKIEREGESQRYKPFROLHNRLLW 860

Qy     444 GYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQ 503
              ||||||||||| |:||:||||:||||||||:||||||||||||||||||||||::: ||
Db     861 HGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANY------CHTSQGDPI 915

Qy     504 GQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEV 538
              | |||||||||||||||||||||||||||||||:|
Db     916 GLIMLGEVALGNMYELKHAS-HISKLPGKHSVKGLGKTTPDPSASITLE----GVEVPL 970

Qy     971 GTGIP-SGVNDTALLLYNEYIVYDIAQVNLKYLLKL 1004
```

Qy   44 KAIPAEKRI-----IRVDPTCPLSSNPGTQVYED----YNCTLNQTNIENNNNKFYIIQL 94
Db   500 KSMPVSRIFKVKDGLAVDPDSGLEDI-AHVVDSNNKYSVVLGLTDIQRNKNSYKVQL 557

Qy   95 LQ-DSNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVS 152

Query Match   22.4%;   Score 638;   DB 1;   Length 994;
Best Local Similarity   32.2%;   Pred. No. 1.8e-36;
Matches 168;   Conservative 100;   Mismatches 195;   Indels 58;   Gaps 19;

```
Db  558 LKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVK 617
Qy  153 HPGKYTLIEVQAEDEAQEAVKVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNT 212
Db  618 RTGRMYPIEIQVDDD--QKLVKHE------SHFFTSKLEISVQNLIKLIEFIDSMNKT 667
Qy  213 MALMDLDVKKMPLGKLSKQQIARGFEALEALEALKGPTDGGQSLEELSSHFYTVIPHNF 272
Db  668 LMEFHIDMDKMPLGKLSAHQIQSAYRVKEIYNVLECGSNTAK-LIDATNRFYTLIPHNF 726
Qy  273 GHSQPPINSPELLQAKKDMLLVLADIELAQQALQAVSEQEKTVEEVPHPLDRDYQLLKCQ 332
Db  727 GVOLPLIETHQQIEDLRQMLDSLAEIEVAYSI----IKSEDVSDACNPLDNHYAQIKTQ 782
Qy  333 LQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKIGNRKLL 389
Db  783 LVALDKNSEEFSILSQVKNTHASTHKSYDLKIVDVFKVSRQGEARRFKPFKLHNRKLL 842
Qy  390 WHGTNMAVVAALLTSGLRI------MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY 445
Db  843 WHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFADMVSKSANYCCTSQ--QNSTGL 900
Qy  446 MFLGEVALG------REHHIN--TDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQ 497
Db  901 MLLSEVALGDMMECTSAKYINKLSNNK------HSCFGRGRTMPDDPTK-SYIRSDG- 949
Qy  498 QVVVPQGGPVPCPEFSSSSTFSQSEYLIIYQESQCRLRYLLEV 538
Db  950 -VEIPYGETI-TDEHLKSSLLYNEYIVVDVAQVNIQILFRM 988
```

RESULT 12
T51353
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) [imported] - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C;Accession: T51353
R;Doucet-Chabeaud, G.; Kazmaier, M.
submitted to the EMBL Data Library, December 1998
R;Reference number: 225379
A;Accession: T51353
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-983 <DOU>
A;Cross-references: EMBL:AJ131705; PIDN:CAA10482.1
A;Experimental source: cultivar landsberg erecta
C;Genetics:
A;Gene: parp-1
C;Function:
A;Description: ADP-ribose polymer synthesis
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match   21.0%;   Score 598.5;   DB 2;   Length 983;
Best Local Similarity   31.1%;   Pred. No. 1e-33;
Matches   171;   Conservative   101;   Mismatches   199;   Indels   79;   Gaps   23;

```
Qy   24 KKGRQAGREEDPPRSTAEALKAIPAEKRIRVDPTCPLSSNPGTQ----VYED---YNC 75
Db  472 KKQRKLPFDKYKIEDTSESLVTVKVKGR------SAVHEASGLQEHCHILEDGNSIYNT 524
Qy   76 TLNQTNIENNNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD 131
Db  525 TLSMSDLSTGINSYYILQIIQE-DKGSDCYVFRKWGRVG-----SDAVHE 582
Qy  132 FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVKVDRGPVRTVTKRVQ 188
Db  583 FKRLFLEKTGNTWESWEQKTNFQKOPGKFLPLD----IDYGVNKQVAKK-E 628
Qy  189 P-----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 244
```

```
Db  629 PFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ 688
Qy  245 EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIE 300
Db  689 RLLTESDPQPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIE 744
Qy  301 LAQAQLQAVSEQEKTVEEVPHPLDRDYQLLKCQLOLLLDSGAPEYKVIQTYLEQTGSNHRCP 360
Db  745 IASRI--VGFDVDSTES----LDDKYKKLLHCDISPLPHDSEDYRLIEKYLNTT 794
Qy  361 T------LQHIWKVNQEGEEDRFQAH-SKIGNRKLLWHGTNMAVVAALTSGLRI---- 408
Db  795 THTEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNGIRIAPPE 854
Qy  409 MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS 468
Db  855 APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGEIHEL-TRAKYMDK 911
Qy  469 PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGGPVPCPEFSSSSTFSQSEYLIIYQES 528
Db  912 PPRGHSTRGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEIVIVDPA 967
Qy  529 QCRLRYLLEV 538
Db  968 QVKLQFLIKV 977
```

RESULT 13
C84719
probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
A;Accession: C84719
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: C84719
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1009 <STO>
A;Cross-references: GB:AE002093; NID:g4432827; PIDN:AAD20677.1; GSPDB:GN00139
A;Gene: At2g31320
C;Genetics:
A;Gene: At2g31320
A;Map position: 2
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match   21.0%;   Score 598.5;   DB 2;   Length 1009;
Best Local Similarity   31.1%;   Pred. No. 1.1e-33;
Matches   171;   Conservative   101;   Mismatches   199;   Indels   79;   Gaps   23;

```
Qy   24 KKGRQAGREEDPPRSTAEALKAIPAEKRIRVDPTCPLSSNPGTQ----VYED---YNC 75
Db  498 KKQRKLPFDKYKIEDTSESLVTVKVKGR------SAVHEASGLQEHCHILEDGNSIYNT 550
Qy   76 TLNQTNIENNNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD 131
Db  551 TLSMSDLSTGINSYYILQIIQE-DKGSDCYVFRKWGRVGNEKIGGNKVEEMSK-SDAVHE 608
Qy  132 FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVKVDRGPVRTVTKRVQ 188
Db  609 FKRLFLEKTGNTWESWEQKTNFQKOPGKFLPLD------------IDYGVNKQVAKK-E 654
Qy  189 P-----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 244
Db  655 PFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ 714
Qy  245 EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIE 300
Db  715 RLLTESDPQPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIE 770
```

RESULT 14

T03657
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) 2 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
C;Accession: T03657
R;Babiychuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Monta
submitted to the EMBL Data Library, November 1997
A;Description: Higher plants possess two poly(ADP-ribose) polymerases.
A;Reference number: Z14992
A;Accession: T03657
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-969 <BAB>
A;Cross-references: EMBL:AJ222589; PIDN:CAA10889.1
C;Genetics:
A;Gene: PARP2
A;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match   20.3%; Score 578; DB 2; Length 969;
Best Local Similarity 30.6%; Pred. No. 2.8e-32;
Matches 166; Conservative 96; Mismatches 195; Indels 86; Gaps 21;

```
Qy  301 LAQAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP 360
Db  771 IASRI--VGEPDVDSTES----LDDKYKKLHCDISPLPHDSEDYRLIEKYLNTT--HAP 820

Qy  361 T------LQHIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRI----- 408
Db  821 THTEWSLELEVFALEREGEPDKYAPHREKLGNKMLLWHGSRLTNEVGILNQGLRIAPPE 880

Qy  409 MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS 468
Db  881 APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGEIHEL-TKAKYMDK 937

Qy  469 PPGFEDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES 528
Db  938 PPRGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEYIVYDTA 993

Qy  529 QCRLRYLLEV 538
Db  994 QVKLQFLKV 1003
```

```
Db  791 EVFSLDRDGELNKYSRYXNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMEGK 850
Qy  419 GIYFASENSKSAGYVIGMKCGA---HHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDS 475
Db  851 GLYFADLVSKSAQY-----CYVDRNNPVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS 904
Qy  476 VIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYL 535
Db  905 TKGLGKTVP---LESEFVKWRDDVVVPCGKPVP-SSIRSSELMYNEYIVYNTSQVKMQFL 960
Qy  536 LEV 538
Db  961 LKV 963
```

RESULT 15

T18600
hypothetical protein AC8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
A;Reference number: Z18996
submitted to the EMBL Data Library, November 1996
R;McMurray, A.
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-727 <WIL>
A;Cross-references: EMBL:Z83097; PIDN:CAB05448.1; GSPDB:GN000023; CESP:AC8.1
A;Experimental source: clone AC8
C;Genetics:
A;Gene: CESP:AC8.1
A;Map position: 5
A;Introns: 31/3; 271/3; 312/3; 348/1; 466/3; 545/3; 617/3; 702/2
A;Accession: T18600

Query Match   18.7%; Score 534; DB 2; Length 727;
Best Local Similarity 28.6%; Pred. No. 2.2e-29;
Matches 162; Conservative 104; Mismatches 211; Indels 90; Gaps 20;

```
Qy   37 RSTPAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQ 96
Db  182 KAIGEYVEALAAKGG--STEPATPASASP-TPPEAETPVLSAEGSPESSNKRPASAEIIE 238

Qy   97 ----DSNRFF--TCWNRWGRVGEVGQSKI--------------------NHFTLRLED 127
Db  239 IDGEGNPDENDFAKKKRRMKKEARLMEVQKKRMKKQSDLLWEYRQIFERMPYTDNISILRE 298

Qy  128 AKKDF-------EKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGP 179
Db  299 NEQDIPEGHDRIAQDVFHEKTKNDWIYRKHFRKMPGMFSYVETDYSE----FVGTNNGH 353

Qy  180 VRTVT--KRVQPCS---LDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA 234
Db  354 KKKITPGSKITPGSKTLLPKSVKEVVMSIFDVENMKSALKSFEIDVNKMPLGRLSHNQIN 413

Qy  235 RGFEALEALEEAL-KGPTDGGQSLEELSSHFYTVIIPHNFGHSQPPPINSPELLQQAKKDML 293
Db  414 LAFEVLNDISDLLVKLPIDASKIL-DFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNML 472

Qy  294 LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 353
Db  473 NALLDIKPAYDQISGGDVPASTSLSIDPVDINVRKLKCIMEPLQQGCDDWNMIHQYLKNT 532

Qy  354 -GSNHRCPT-LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH 411
Db  533 HGATHDLKVELIDILKVNRDNESSKFKRH--IGNRRLLWHGSGKMNFAGILSQGLRIAPP 590

Qy  412 ----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGR-EHHINTDNPSL 466
Db  591 EAPVSGYMFGKGVYFADMFSKSFFY---CRANAKEEAYLLCDVALGNVQQLMASKNVSR 647

Qy  467 KSPPPGFEDSVIARG--------HTEPD----PTQDTELELDGQQVVVPQGQPVPCPEFS 513
```

```
             :: | ||| ||   |              : | ||   |    :: | ||: |
Db   648 QTLPAGFQSVGVGVGRQCPREIGSYIKNPDGYTVPLGITYMQLQGKQNV------- 694

Qy   514 SSTFSQSEYLIYQESQCRLRYLLEVHL 540
             :|:::|   | :|:||: |  :
Db   695 DYHLLYNEFIVYDVDQIQLKYLVRVKM 721
```

Search completed: August 29, 2002, 07:59:33
Job time: 280 sec

OM protein - protein search, using sw model

```
          GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.
```

Run on: August 29, 2002, 07:57:16 ; Search time 124.84 Seconds
                                         (without alignments)
                                         474.226 Million cell updates/sec

Title: US-09-701-586B-8
Perfect score: 2813
Sequence: 1 MAPKRKASVQTEGSKKQRQG.........EYLIYKESQCRLRYLLEIHL 533

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   A_Geneseq_032802:*
        1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
        2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
        3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
        4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
        5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
        6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
        7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
        8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
        9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
       10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
       11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
       12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
       13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
       14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
       15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
       16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
       17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
       18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
       19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
       20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
       21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
       22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*
```

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 2813 | 100.0 | 533 | 21 | AAY51177 | Murine PARP1 (long |
| 2 | 2776.5 | 98.7 | 528 | 21 | AAY51178 | Murine PARP1 (shor |
| 3 | 2241 | 79.7 | 533 | 21 | AAY51175 | Human brain PARP3 |
| 4 | 2241 | 79.7 | 540 | 21 | AAY51176 | Human uterus type |
| 5 | 2229 | 79.2 | 533 | 22 | AAU29021 | Human PARP-3 prote |
| 6 | 717 | 25.5 | 653 | 21 | AAY68834 | A poly(ADP-ribose) |
| 7 | 705 | 25.1 | 1013 | 17 | AAR99642 | Poly(ADP-ribose) |
| 8 | 702.5 | 25.0 | 522 | 22 | AAU29022 | Mouse PARP-2 prote |
| 9 | 700 | 24.9 | 1014 | 21 | AAY58043 | Human poly (ADP-ri |
| 10 | 700 | 24.9 | 1014 | 22 | AAU29019 | Human PARP-1 prote |
| 11 | 700 | 24.9 | 1014 | 22 | AAB66296 | Human tankyrase2 r |
| 12 | 699 | 24.8 | 1014 | 20 | AAY33699 | Human poly(ADP-rib |
| 13 | 698 | 24.8 | 521 | 22 | AAB60693 | Human poly(ADP-rib |
| 14 | 698 | 24.8 | 570 | 21 | AAY51174 | Human brain PARP2 |
| 15 | 698 | 24.8 | 570 | 22 | AAB11480 | Human brain poly-A |
| 16 | 696.5 | 24.8 | 534 | 21 | AAB42909 | Human ORFX ORF2673 |
| 17 | 696.5 | 24.8 | 534 | 22 | AAU29023 | hPARP2. Homo sapi |
| 18 | 696.5 | 24.8 | 583 | 22 | AAB47029 | Human nuclear NAD+ |
| 19 | 696 | 24.7 | 1014 | 21 | AAY49939 | Human nuclear NAD+ |
| 20 | 695.5 | 24.7 | 534 | 21 | AAU29020 | Human PARP-2 prote |
| 21 | 690 | 24.5 | 637 | 21 | AAY68835 | The poly(ADP-ribos |
| 22 | 646.5 | 23.0 | 531 | 22 | AAB93513 | Human protein sequ |
| 23 | 619 | 22.0 | 557 | 22 | ABB64431 | Drosophila melanog |
| 24 | 599.5 | 21.3 | 1063 | 22 | AAB47032 | Fusion protein PAR |
| 25 | 584.5 | 20.8 | 969 | 21 | AAY68833 | A poly(ADP-ribose) |
| 26 | 584.5 | 20.8 | 980 | 21 | AAY68839 | A poly(ADP-ribose) |
| 27 | 491 | 17.5 | 982 | 20 | AAY28464 | Maize poly ADP-rib |
| 28 | 456 | 16.2 | 379 | 22 | AAU21687 | Novel human neopla |
| 29 | 435 | 15.5 | 360 | 22 | AAB47030 | N-terminal fragmen |
| 30 | 398 | 14.1 | 1010 | 21 | AAY68840 | Fusion protein of |
| 31 | 378 | 13.4 | 294 | 22 | AAU20129 | Human DNA repair a |
| 32 | 378 | 13.4 | 294 | 22 | AAU21810 | Novel human neopla |
| 33 | 364.5 | 13.0 | 287 | 22 | AAB47031 | C-terminal fragmen |
| 34 | 275 | 9.8 | 1730 | 22 | AAU33242 | Novel human secret |
| 35 | 270.5 | 9.6 | 1724 | 21 | AAY54373 | cDNA sequence enco |
| 36 | 270.5 | 9.6 | 1724 | 22 | AAB51022 | Human minor vault |
| 37 | 261 | 9.3 | 1099 | 22 | AAB66301 | Human tankyrase2 e |
| 38 | 246.5 | 8.8 | 227 | 22 | AAU20130 | Human DNA repair a |
| 39 | 246.5 | 8.8 | 227 | 21 | AAU21811 | Novel human neopla |
| 40 | 203.5 | 7.2 | 190 | 22 | AAU21688 | Novel human neopla |
| 41 | 156.5 | 5.6 | 1181 | 22 | ABB60894 | Drosophila melanog |
| 42 | 156.5 | 5.6 | 1181 | 22 | AAB66297 | Drosophila tankyra |
| 43 | 146 | 5.2 | 1327 | 22 | AAB27212 | Human tankyrase I |
| 44 | 146 | 5.2 | 1327 | 21 | AAY44402 | Human tankyrase. |
| 45 | 146 | 5.2 | 1327 | 22 | AAB66279 | Human tankyrase1 s |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## ALIGNMENTS

```
RESULT  1
AAY51177
ID   AAY51177 standard; Protein; 533 AA.
XX
AC   AAY51177;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Murine PARP1 (long) homologue protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP1.
XX
OS   Mus sp.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;  99WO-EP03889.
XX
PR   05-JUN-1998;  98DE-1025213.
PR   01-MAR-1999;  99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M,  Hoeger T,  Kroeger B,  Otterbach B,  Lubisch W,  Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AZ244290.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
```

```
PT  conditions -
XX
PS  Claim 4; Page 67-69; 96pp; German.
XX
CC  This invention describes novel human and murine poly(ADP-ribose)
CC  polymerase (PARP) homologues, which are characterised by an amino acid
CC  sequence with a functional NAD^+-binding site and no zinc finger
CC  sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC  sequences, PARP homologues and antibodies are useful for analytic
CC  detection of PARP homologues and for identifying PARP effectors or
CC  binding partners, as well as for determining their effectiveness.
CC  PARP-binding partners are useful for the diagnosis or therapy of a
CC  disease condition, which is the result of a PARP protein, especially an
CC  energy deficiency, which may comprise tissue damage from cell death
CC  following necrosis or apoptosis. The disease condition may be chosen
CC  from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC  in particular neurotoxic disturbances, etc. This sequence represents the
CC  murine PARP1 protein used in the method of the invention.
XX
SQ  Sequence  533 AA;

Query Match             100.0%;  Score 2813;  DB 21;  Length 533;
Best Local Similarity   100.0%;  Pred. No. 3e-243;
Matches  533;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1  MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  mapkrkasvqtegskkqrqgteeedsfrstaealraapadnrvirvdpscpfsrnpgiqv  60

Qy   61  HEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLEDAKK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  hedydctlnqtnignnnnkfyiiqlleegsrffcwnrwgrvgevggskmnhftcledakk 120

Qy  121  DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVK 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  dfkkkfwektknkweerdrfvaqpnkytlievqgeaesqeavvkalspqvdsgpvrtvvk 180

Qy  181  PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  pcsldpatqnlitnifskemfknamtlmnldvkkmplgkltkqqiargfealealeeamk 240

Qy  241  NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  nptgdgqsleelsscfytviphnfgrsrpppinspdvlqakkdmllvladielaqtlqaa 300

Qy  301  PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWK 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  pgeeeekveevphpldrdyqllrcqlqlldsgeseykaiqtylkqtgnsyrcpnlrhvwk 360

Qy  361  VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  vnregegdrfqahsklgnrrllwhgtnvavvaailtsglrimphsggrvgkgiyfasens 420

Qy  421  KSAGYVTTMHCGGHQVGYMFIGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  ksagyvttmhcgghqvgymfigevalgkehhitiddpslkspppgfdsviargqtepdpa 480

Qy  481  QDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 533
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  qdieleldgqpvvvpqgppvvqcpsfksssfsqseyliykesqcrlrylleihl 533

RESULT  2
AAY51178
ID   AAY51178 standard; Protein; 528 AA.
XX
AC   AAY51178;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Murine PARP1 (short) homologue protein.
XX
```

```
KW  PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW  diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW  ischemic tissue damage; PARP1.
XX
OS  Mus sp.
XX
PN  WO996572-A2.
XX
PD  16-DEC-1999.
XX
PF  04-JUN-1999;  99WO-EP03889.
XX
PR  05-JUN-1998;  98DE-1025213.
PR  01-MAR-1999;  99DE-1008837.
XX
PA  (BADI ) BASF AG.
XX
PI  Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR  WPI; 2000-087218/07.
DR  N-PSDB; AA244291.
XX
PT  Novel genes and proteins, antibodies and binding partners useful in
PT  diagnosis and therapy of energy deficiency associated disease
PT  conditions -
XX
PS  Claim 4; Page 71-73; 96pp; German.
XX
CC  This invention describes novel human and murine poly(ADP-ribose)
CC  polymerase (PARP) homologues, which are characterised by an amino acid
CC  sequence with a functional NAD^+-binding site and no zinc finger
CC  sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC  sequences, PARP homologues and antibodies are useful for analytic
CC  detection of PARP homologues and for identifying PARP effectors or
CC  binding partners, as well as for determining their effectiveness.
CC  PARP-binding partners are useful for the diagnosis or therapy of a
CC  disease condition, which is the result of a PARP protein, especially an
CC  energy deficiency, which may comprise tissue damage from cell death
CC  following necrosis or apoptosis. The disease condition may be chosen
CC  from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC  in particular neurotoxic disturbances, etc. This sequence represents the
CC  murine PARP1 protein used in the method of the invention.
XX
SQ  Sequence  528 AA;

Query Match             98.7%;  Score 2776.5;  DB 21;  Length 528;
Best Local Similarity   99.1%;  Pred. No. 5.5e-240;
Matches  528;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy    1  MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  mapkrkasvqtegskkqrqgteeedsfrstaealraapadnrvirvdpscpfsrnpgiqv  60

Qy   61  HEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLEDAKK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  hedydctlnqtnignnnnkfyiiqlleegsrffcwnrwgrvgevggskmnhftcledakk 120

Qy  121  DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVRTVVK 180
         |||||||||||||||||||||||||||||||||||||||||         |||||||
Db  121  dfkkkfwektknkweerdrfvaqpnkytlievqgeaesqeavk----vdsgpvrtvvk 175

Qy  181  PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  176  pcsldpatqnliitnifskemfknamtlmnldvkkmplgkltkqqiargfealealeeamk 235

Qy  241  NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  236  nptgdgqsleelsscfytviphnfgrsrpppinspdvlqakkdmllvladielaqtlqaa 295

Qy  301  PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWK 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  296  pgeeeekveevphpldrdyqllrcqlqlldsgeseykaiqtylkqtgnsyrcpnlrhvwk 355
```

```
Qy      1 MAPKRASVQTEG--SKKQRQGTEEEDSFRSTAEALRAAPADNRVIRRVDPSCPFSRNPGI  58
          |||| |||||||    ||| ||  |||||||:|| ::||:||||:||:|:||  ||||||
Db      1 mapkpkpwvqtegpekkkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt  60

Qy     59 QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED  117
          ||:||:||||||||||||||||:||||||||||::-||||||||||||||::||||||||
Db     61 qyyedynctlnqtniennnnkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftried  120

Qy    118 AKKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT  177
          |||||:|||:|| |:|| :: ||  |:|||||: | |:|:||||  ----vdrgpvrt  175
Db    121 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvk----vdrgpvrt  175

Qy    178 V---VKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA  234
          |    :|||||||||||||:|||| :||| |:|||||:|||||:|:|:|||:|||||||
Db    176 vtkrvqpcsldpatqkliitnifskemfkntmalmdldvkkmpigkslskqqiargfealea  235

Qy    235 LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELA  294
          ||||| | ||||| |||||||:| |||||:|||||| |||:|||||||||||||||||
Db    236 leealkptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladiela  295

Qy    295 QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPN  354
          |  ||| :: ||||||| |||||||:|||||||||||| |||||||:|| |||:: |||
Db    296 qalqav-seqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcpt  354

Qy    355 LRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIY  414
          |:|:|||||||||||||||||||||||||||||||||||:|||||||||||| |||| |
Db    355 lqhiwkvnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiy  414

Qy    415 FASENSKSAGYYTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQ  474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db    415 fasensksagyyvtmhcgahvgymflgevaigrehhintdnpslksppgfdsviargh  474

Qy    475 TEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL  533
          |||||:||:||||||||||| || || || ||::||:|:|||:||||:|||||||||:|
Db    475 tepdptqdteleldgqqvvvpqgqpvvpcpefssstfsqseyllyqesqcrlrylllevhl  533
```

RESULT 3

```
AAY51175
ID   AAY51175 standard; Protein; 533 AA.
XX
AC   AAY51175;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human brain PARP3 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP3.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;  99WO-EP03889.
XX
PR   05-JUN-1998;  98DE-1025213.
PR   01-MAR-1999;  99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AA244288.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
PS   Claim 4; Page 57-59; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula Cx_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP3 protein used in the method of the invention.    .
XX
SQ   Sequence 533 AA;
```

```
Query Match             79.7%;  Score 2241;  DB 21;  Length 533;
Best Local Similarity   80.1%;  Pred. No. 5.8e-192;
Matches  432;  Conservative  37;  Mismatches  58;  Indels  12;  Gaps   5;
```

```
Qy    361 VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENS  420
          ||||||||||||||||||||||||||||||||||||:|||||||||||| |||| ||||
Db    356 vnregedrfqahsklgnrrllwhgtnvavvaailtsglrimphsggrvgkgiyfasens  415

Qy    421 KSAGYYTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db    416 ksagyyttmhcgghqvgymflgevalgkehhitiddpslksppgfdsviargqtepdpa  475

Qy    481 QDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL  533
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    476 qdieleldgqpvvvpqgppvvqcpsfkssssfsqseyllykesqcrlrylleihl  528
```

RESULT 4

```
AAY51176
ID   AAY51176 standard; Protein; 540 AA.
XX
AC   AAY51176;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human uterus type 2 PARP3 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP3.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;  99WO-EP03889.
XX
PR   05-JUN-1998;  98DE-1025213.
PR   01-MAR-1999;  99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AA244289.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
```

```
XX
PS  Claim 4; Page 62-64; 96pp; German.
XX
CC  This invention describes novel human and murine poly(ADP-ribose)
CC  polymerase (PARP) homologues, which are characterised by an amino acid
CC  sequence with a functional NAD^+-binding site and no zinc finger
CC  sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC  sequences, PARP homologues and antibodies are useful for analytic
CC  detection of PARP homologues and for identifying PARP effectors or
CC  binding partners, as well as for determining their effectiveness.
CC  PARP-binding partners are useful for the diagnosis or therapy of a
CC  disease condition, which is the result of a PARP protein, especially an
CC  energy deficiency, which may comprise tissue damage from cell death
CC  following necrosis or apoptosis. The disease condition may be chosen
CC  from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC  in particular neurotoxic disturbances, etc. This sequence represents the
CC  human PARP3 protein used in the method of the invention.
XX
SQ  Sequence  540 AA;
XX

Query Match        79.7%;   Score 2241;  DB 21;  Length 540;
Best Local Similarity  80.1%;  Pred. No. 6e-192;
Matches  432;  Conservative  37;  Mismatches  58;  Indels  12;  Gaps  5;

Qy      1 MAPKRKASVQTEG--SKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58
          ||| ||| ||||||  ||| ||:|| |||:||||||||||||||||||:|| |||||
Db      8 mapkpkpwvqtegpekkkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt 67

Qy     59 QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117
          || ||||||||| :|||||||||||||||| :: ||  ||||||||||||| :|||||||
Db     68 qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftrled 127

Qy    118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
          ||||| ||| |||||:| |||| :|| ||||||||:|| :|||||||||:-----||||||
Db    128 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvk-----vdrgpvrt 182

Qy    178 V---VKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA 234
          |    : ||||||||:|||||||||||||: ||| |ld|||||:|| |||||||||||||
Db    183 vtkrvqpcsldpatqklitnifskemfkntmalmdldvkkmpigklskqqiargfealea 242

Qy    235 LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELA 294
          ||||:|| |||||||||||| |||||||||||: ||||||||| ||||||||||||||||
Db    243 leealkgptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladiela 302

Qy    295 QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPN 354
          | ||  |||:|  ||||||||||||||:|||||||||:::||| |||:|:||| :::|||
Db    303 qalqav-seqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcpt 361

Qy    355 LRHYWKVNREGEGDRFQAHSKLGNRRLIWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIY 414
          |:|::||| :||||||||||:||||: ||||||:|||||||||||||||||||||||:||||
Db    362 lqhiwkvnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiy 421

Qy    415 FASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQ 474
          |||||||::||| | |||:||| |||||||||| :|||: |||||:|||| ||||||||
Db    422 fasenskssagyvymkcgahhvqymflgevalgrehhintdnpslksplpqfdsviargh 481

Qy    475 TEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 533
          |||| |  |||:|||| ||||||  ||| | ::||::|||:|| :||||||||||||:||
Db    482 tepdtqdteleldgqqvvvpqggqvvpqgpvpcpefssstfsqseylyqesqcrlryllevhl 540

RESULT 5
AAU29021
ID  AAU29021 standard; Protein; 533 AA.
XX
AC  AAU29021;
XX
DT  18-DEC-2001  (first entry)
XX
DE  Human PARP-3 protein.
XX
KW  Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
```

```
KW  cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW  immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW  oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW  meningitis-associated intracranial complication; ischaemia;
KW  inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS  Homo sapiens.
XX
PN  WO20016495S-A1.
XX
PD  07-SEP-2001.
XX
PF  01-MAR-2001; 2001WO-US06572.
XX
PR  02-MAR-2000; 2000US-0517467.
XX
PA  (ISIS-) ISIS PHARM INC.
XX
PI  Popoff I, Cowsert LM;
XX
DR  WPI; 2001-602570/68.
DR  N-PSDB; AAS45590.
XX
PT  Antisense compound useful for treating hyperproliferative,
PT  neurological, inflammatory and autoimmune disorders and diabetes
PT  inhibits human PARP -
XX
PS  Example 13; Page 105-107; 168pp; English.
XX
CC  The invention relates to antisense oligonucleotides targeted to human
CC  PARP nucleic acid and inhibiting expression of human PARP. PARP
CC  (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC  decondensation, DNA replication, DNA repair, gene expression, malignant
CC  transformation, cellular differentiation and apoptosis. The antisense
CC  oligonucleotide inhibitors are useful for inhibiting the expression of
CC  PARP in human cells or tissues. They are also useful for treating a
CC  human with a disease associated with PARP especially hyperproliferative
CC  disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC  neurological (e.g parkinsonism, meningitis-associated intracranial
CC  complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC  arthritis) and diabetes. The present sequence is a PARP protein,
CC  the cdNA encoding which was used to design the antisense
CC  oligonucleotides.
XX
SQ  Sequence  533 AA;

Query Match         79.2%;    Score 2229;  DB 22;  Length 533;
Best Local Similarity  79.8%;   Pred. No. 7e-191;
Matches  430;   Conservative  37;  Mismatches  60;  Indels  12;  Gaps  5;

Qy      1 MAPKRKASVQTEG--SKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58
          ||| ||| ||||||  ||| |||:|| |||:|||||||||||||||||:|| |||||
Db      1 mapkpkpwvqtegpekkkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt 60

Qy     59 QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117
          || ||||||||| :||||||||||||||| :: ||  |||||||||||||| :|||||||
Db     61 qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftrled 120

Qy    118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
          ||||| ||| |||||:| |||| :|| ||||||||:|| :||||||||-----||| |||
Db    121 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvk-----vdrapvrt 175

Qy    178 V---VKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA 234
          |    : ||||||||:||||||||||||: ||| |ld|||||:|| |||||||||||||
Db    176 vtkrvqpcsldpatqklitnifskemfkntmalmdldvkkmpigklskqqiargfealea 235

Qy    235 LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELA 294
          ||||:|| |||||||||||| |||||||||||: |||||||||||||||||||||||||
Db    236 leealkgptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladiela 295

Qy    295 QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPN 354
          | ||  |||:|  ||||||||||||||:|||||||||::||| |||:|:||| :::|||
```

XX
SQ   Sequence   653 AA;

Query Match   25.5%;   Score 717;   DB 21;   Length 653;
Best Local Similarity   34.9%;   Pred. No. 3.7e-55;
Matches 190;   Conservative 91;   Mismatches 205;   Indels 58;   Gaps 17;

```
Qy     13 GSKKQRQGTEEEDSFRS---TAEALRAAPADNRV---IRVDPSCPFSRNPGIQVHED-Y  64
          :|| :|:|||::||||||||::|||||:||||||||    ||| ||  :: :  |:
Db    138 gakevlkggdeevevkkekmvtatkkgaavldqhipdhikkvnyhv------lqvgdely 190

Qy     65 DCTLNQTNIGNNNNKFYIIQLLEE--GSRRFCWNRRGRVGEVGQSKMNHFTCLED-AKKD 121
          |||||||||:|||||||||||||| || |||||||||:||::|:|| || ||  |||
Db    191 datlnqtnvgdnnnkfyilqvlesdaggsfmvynrwgrvgvrgqdklhgpsptrdqaiye 250

Qy    122 FKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKALSPQVDSGPVRTVVK 180
          |::| |::::|::||||:||:|||||:|||:|| ||:|| ::|:|:    ||  :|
Db    251 fegkfhnktnnhwsdrknfkcyakkytwlemdygetek------elekgsitdqik 300

Qy    181 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 240
          |:||:|:| |||:||||:::|:|||:|||:|: :|::|:||::|:| |::|::||||||
Db    301 etkletriaqfisilcnismkqrmveigynaekiplgkirkatilkgyhvlkrisdvis 360

Qy    241 NPTGDGQSLEELSSCFYTVIPHNFG--RSRPPPINSPDVLQAKKDMLLVLADIELAQTLQ 298
          |  ::| :||||| :| :|:|| | :|  ||  :: | |:|:||||| :|:  |
Db    361 k--adrrheqltgefytviphdfgfrkmrefiidtpqklkaklemvealgeieiatkll 418

Qy    299 AAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT--GNSYRCPNL 355
          :|  : ::||||::   |:|::|  |: ::|||:|::||:|:|||:::  |:::|||:
Db    419 ------eddssdqddplyarykqlhcdftpleadsdeysmiksyirnthgkthssytvdi 472

Qy    356 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK 411
          ||:|||:||||||||||||| ||||||||||||:|||||||:|||:||    |||||||
Db    473 vqifkvsrhgeterfqkfastrnrmllwhgsrlsnwagilsqgiriappeapvtgymfgk 532

Qy    412 GIYFASENSKSAGYVTMHCGGHQV---GYMFLGEVALGKEHHITIDDPSLKSPPPGFDS 468
          || ||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    533 gvyfadmfsksany------cyaseacrsgvllcevalgdmnellnadydannlpkgklr 587

Qy    469 VIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYL 528
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    588 skgvyqtapnm---veskvaddgvvvplgepkqeps-krggllyneyivynvdqirmryv 643

Qy    529 LEIH 532
          |:||
Db    644 lhvn 647
```

RESULT   7
AAR99642
ID   AAR99642 standard; Protein; 1013 AA.
XX
AC   AAR99642;
XX
DT   31-OCT-1996   (first entry)
XX
DE   Poly(ADP-ribose) polymerase contg. DNA-binding domain.
XX
KW   PARP; DNA-binding domain; poly(ADP-ribose) polymerase; gene therapy;
KW   tumour treatment; DNA repair; over-expression.
XX
OS   Homo sapiens.
XX
PN   WO9618737-A2.
XX
PD   20-JUN-1996.
XX
PF   15-DEC-1995;   95WO-DE01817.
XX
PR   16-DEC-1994;   94DE-444949.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

```
Db    296 qalqav-seqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcpt 354

Qy    355 LRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIY 414
          |:|||||:||||||||||||| ||||||||||||||||||||||||||:|||||||||||
Db    355 lqhiwkvnqegeedrfqahskignrkllwhgtnmavvaailtsglrimphsggrvgkgiy 414

Qy    415 FASENSKSAGYVTMHCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQ 474
          ||||||||||||||||| ||||||||||||||||||||||||:|||||||||||||||
Db    415 fasensksagyyvigmkcgahhvgymflgevalgrehhintdnpslkspppgfdsviargh 474

Qy    475 TEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|:
Db    475 tepdtqdteleldgqqvvvpqgppvpcpefssstfsqseyliyqesqcrlryllevhl 533
```

RESULT   6
AAY68834
ID   AAY68834 standard; Protein; 653 AA.
XX
AC   AAY68834;
XX
DT   16-MAY-2000   (first entry)
XX
DE   A poly(ADP-ribose) polymerase NAP protein of Zea mays.
XX
KW   NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
KW   programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
KW   pest; drought; heat; fungi; nematode; seed shatter.
XX
OS   Zea mays.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference   1..159
FT   /note= "these residues are specifically claimed in
FT   claim 18"
XX
PN   WO200004173-A1.
XX
PD   27-JAN-2000.
XX
PF   12-JUL-1999;   99WO-EP04940.
XX
PR   17-JUL-1998;   98US-0118276.
XX
PA   (PLBZ) PLANT GENETIC SYSTEMS NV.
XX
PI   Bablychuk E, Kushnir S, De Block M;
XX
DR   WPI; 2000-182436/16.
DR   N-PSDB; AA260616.
XX
PT   Modulating cell death, growth and stress resistance in eukaryotes,
PT   specifically plants, used, e.g. to impart fungus or nematode resistance
PT   -
XX
PS   Claim 18; Page 92-95; 126pp; English.
XX
CC   The present sequence represents the NAP protein of Zea mays. This
CC   protein is a poly(ADP-ribose) polymerase (PARP) protein (also known as
CC   poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
CC   cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide
CC   sequences can be used for modulation of programmed cell death in
CC   eukaryotic cells. The method is used, specifically in plants, to induce,
CC   or protect against, programmed cell death, depending on the extent to
CC   which PARP activity is reduced. Reducing expression of endogenous NAP
CC   class PARP only is also used to modulate programmed cell death, to
CC   increase growth rate and to produce plant cells that are more tolerant
CC   of stress (cold, chemical treatments, pathogens, pests, drought, heat,
CC   etc., or during transformation). Particular applications are generation
CC   of plants that are resistant to fungi or nematodes; are male or female
CC   sterile; or have better seed-shatter properties. The methods are also
CC   used to improve growth of transformed plant cells (and derived calli or
CC   complete plants).

```
XX
PI   Buerkle A, Kuepper J, Zur Hausen H;
XX
DR   WPI; 1996-300654/30.
DR   N-PSDB; AAT13732.
XX
PT   Vectors contg. insert encoding DNA-binding domain of
PT   poly(ADP-ribose) polymerase - useful for gene therapy, esp. of
PT   tumours
XX
PS   Disclosure; Fig 1; 22pp; German.
XX
CC   The present sequence is that of a poly(ADP-ribose) polymerase (PARP)
CC   contg. a DNA-binding domain (DBD). It is encoded by AAT13732, from
CC   which a fragment comprising nucleotides 67-1220 (-29 to + 1127) encoding
CC   the DBD can be inserted into vectors which are used for gene therapy.
CC   Over-expression of the DBD inhibits the DNA repair function of PARP, so
CC   the vectors are useful for gene therapy or tumours, esp. in combination
CC   with conventional chemo- and/or radiotherapy.
XX
SQ   Sequence  1013 AA;

Query Match             25.1%;   Score 705;  DB 17;  Length 1013;
Best Local Similarity   33.8%;   Pred. No. 8.8e-54;
Matches  183;  Conservative 101;  Mismatches 218;  Indels  40;  Gaps  17;

Qy     1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
         |:|| ||  ||   ::|| ::|      :||      |                |:|:|  |
Db   492 vaprgksga---alskkskgqvkeeginksekrmkltlkggaavdpdsglehsahvlekg 548

Qy    61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
         || : :|  :|||  | |  |||||||     |  || :| :|||||  ||  | | ||
Db   549 gkvfsatlglvdivkgtnsyykiqlleddkenrywifrswgrvgtvigsnkleqmpsked 608

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
         |:|| ||:||| ||   |  :|| |   :|| |||| ::||  :|  | ||:|   || |
Db   609 aiehfmklyeektgnawhskn-ftkypkkfypleid-yggqeeavkkl---tvnpgtksk 663

Qy   178 VVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
         | || | | | | ||   |||::| |  |  |||| |  | :|||:|||:|| || |||
Db   664 lpkp-----vqdlikmifdvesmkamveyeidiqkmpigklskrqiqaaysilsevqq 717

Qy   238 AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTL 297
         | |:   | |    :|| | :: |   ||    :| ||  || |||:|||| :| :|
Db   718 avsqgsdsqil-disnrfytliphdfgmkkpplinnadsvqakvemldnlldievaysl 776

Qy   298 QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGSEYKAIQTYLKQ---TGNSYRCP 353
         ::| |  :|      :| | |:   | :|:|   ::  |  |:|  |       |:|
Db   777 lrggsddssk-----dpidvnyeklktdikvvdrdseeeaeiirkyyvknthatthnaydl- 830

Qy   354 NLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRV 409
         :  :|||  ||   :| ||||||:| ::||||||||||||||||||||||   :|   :
Db   831 evidifkieregecqrykpfkqlhnrrllwhgsrtnfagilsgglriappeapvtgymf 890

Qy   410 GKGIYFASENSKSAGYVTMHCGGHOVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSV 469
         ||||||| :|  ||  | | | |  :  |  || | |||  ||  | |:|  ||||||
Db   891 gkgiyfadmvsksanychtsq-gdpigliilgevalgnmyelk-hashisklpkgkhsv 947

Qy   470 IARGQTEPDPAQDIELELDGOPVVVPQGPVQCPSFKSSSFSQSEYLIYKESQCRLRYLL 529
         || |  | |   |  :||: |  |  |:| ||:::| ||: |  :| :|  : ||||||
Db   948 kglgkttpdpsani--sldg--vdvplgtgis-sgvndtsllyneyivydiaqvnlkyll 1002

Qy   530 EI 531
         :|
Db  1003 kl 1004
```

```
RESULT 8
AAU29022
ID   AAU29022 standard; Protein; 522 AA.
XX
AC   AAU29022;
```

```
XX
DT   18-DEC-2001 (first entry)
XX
DE   Mouse PARP-2 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Mus musculus.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I,  Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45597.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 109-111; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cdNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence  522 AA;

Query Match             25.0%;   Score 702.5;  DB 22;  Length 522;
Best Local Similarity   34.4%;   Pred. No. 5.3e-54;
Matches  186;  Conservative  89;  Mismatches 200;  Indels  65;  Gaps  19;

Qy    20 GTEEEDSFRSTAEALRA------APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTLNQ 70
         |  ||  :   :::|:           |||  |  :| :|   |        :|  |  |
Db    12 ggkdadrtkdnrdsvktlllkgkap------vdpecaaklgkahvycegddvydvmlq 64

Qy    71 TNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFKKK 125
         ||: ||||||:|||:||       |  |||| |  || | ||||     |: |||||||
Db    65 tnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgktgqhsl--vtcsgdlnkakeifqkk 122

Qy   126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLD 185
         | :||||:||||  | |:| :| |:|       |:        |:|:|      |  ||
Db   123 fldktknnwedrenfekvpgkydmlqmdyaastqd-------esktkeeetlkpesqld 174

Qy   186 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 245
         | :| :|   :|    |   |:| | |:|:|| | |||| | |:    ||: |   |||
Db   175 lrvqellkicnvgtmeemiemkydtkrapigkltvaqikagyqslkkiedcir--agg 232
```

```
Qy   246  -GQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEE 304
              |:::|:|  ||||||||:|| | |:|  : ||:||:|  |
Db   233  hgralveacnefytriphdfgslsippvirtekelsdkvkllealgdieialklvks---e 289

Qy   305  EEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------NLRH 357
              |||| |  ||| |:| ::  |:|||:|: : |:| | : ||
Db   290  rqgle--hpldqhyrnlhcalrpldhesnefkvisqylqst---hapthkdytmtlld 342

Qy   358  VWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGI 413
              |:::|  ||||||||:|:   |  ||||||: |
Db   343  vfevekegekeafr--edlpnrmliwhgsrlsnwvgilshglrvappeapltgymfgkgi 400

Qy   414  YFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 473
              || :   ||:|||||||||| |:| || | | |
Db   401  yfadmssksanycfasrl--kntglllsevalgqcnelleanpkagqllrgkhstkgmg 458

Qy   474  QTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSSF--SQSEYLIYKESQCRLRYLLEI 531
              :|:| |: :||:|:| :  ||| | :::|::||: |
Db   459  kmapspahfi--tlngs--tvplgpasdtgilnpegytlnypefivyspnqvrmrylki 514
```

```
RESULT   9
AAY58043
ID   AAY58043 standard; peptide; 1014 AA.
XX
AC   AAY58043;
XX
DT   14-MAR-2000 (first entry)
XX
DE   Human poly (ADP-ribose) polymerase protein sequence.
XX
KW   Poly (ADP-ribose) polymerase; DNA excision repair; DNA damage; PARP;
KW   carcinogen; cancer therapy; base modification; human; cancer; UvrA; UvrB.
XX
OS   Homo sapiens.
XX
PN   WO9958971-A2.
XX
PD   18-NOV-1999.
XX
PF   13-MAY-1999; 99WO-CA00444.
XX
PR   13-MAY-1998; 98US-0078347.
XX
PA   (LEXX/) LE X.
PA   (WEIN/) WEINFELD M.
PA   (XING/) XING J Z.
XX
PI   Le X, Weinfeld M, Xing JZ;
XX
DR   WPI; 2000-062317/05.
XX
PT   Quantifying nucleic acid modifications useful for identifying and
PT   detecting exposure to carcinogens, in early risk assessment for cancer,
PT   and in monitoring cancer therapy -
XX
PS   Example 6; Fig 9; 65pp; English.
XX
CC   This is the human poly (ADP-ribose) polymerase protein (PARP) sequence.
CC   PARP is a 116kD DNA-binding protein that binds tightly to single-strand
CC   and double-strand breaks. The PARP protein and Escherichia coli UvrA and
CC   UvrB proteins (see AAY58041-AY58042) are used in a method for the
CC   detection and quantitation of DNA base damage. The method involves
CC   quantifying at least one modification of interest in a nucleic acid
CC   sequence through the used of capillary electrophoresis, laser induced
CC   fluorescence and a fluorescently labeled polypeptide (e.g. fluorescently
CC   labeled PARP or fluorescently labeled anti-PARP antibody) that is
CC   specific for the modification. The new method is useful for quantifying
CC   nucleic acid modifications, especially a mutation (including deletions,
CC   insertions and substitutions), mismatch, DNA adduct, or strand break
CC   (single or double). The method is useful for identifying and detecting
CC   exposure to carcinogens, and in early risk assessment for cancer, and is
CC   also useful in monitoring cancer therapy.
```

```
XX
SQ   Sequence   1014 AA;

Query Match             24.9%;  Score 700;  DB 21;  Length 1014;
Best Local Similarity   33.6%;  Pred. No. 2.5e-53;
Matches 182;  Conservative 102;  Mismatches 219;  Indels 38;  Gaps 17;

Qy     1  MAPKRKASVQTEGSKKQRGQTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
             :|:|: :                          :: :::   :|
Db   493  vaprgksga--alskkskgqvkeeginksekrmkltlkggaavdpdsglehsahvlekg 549

Qy    61  HEDYDCTLNQTNIGNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
             |   ::|   | :  |:| :: |    :::   ||||| ::|||   | ::|
Db   550  gkvfsatlglvdivkgtnsyykiglleddkenrywifrswgrvgtvigsnkleqmpsked 609

Qy   118  AKKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
             ::  : ||:| :|  |  ||| :  |:   |  :|: ||| ||:||:|: ||| |
Db   610  aleqfmklyeektgnawhskn-ftkypkkfypleid-yggdeeavkkl---tvnpgtksk 664

Qy   178  VVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
             : ||         :|| ::          |:||||||| : :: :::
Db   665  lpkp------vqdlikmifdvesmkkamveyeidlqkmplgkiskrqiqaaysllsevqq 718

Qy   238  AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTL 297
             |: : :|  ||| :||:||||| ||||||:|||||||||||||||||||||||||:|||
Db   719  avsqgssdsqil-dlsnrfytliphdfgmkkpplinnadsvqakvemlnlldievaysl 777

Qy   298  QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR 356
             : :    | :|| : | ||       ::|:  :: || ||:   |
Db   778  lrggsdddssk-----dpidvnyeklktdikvvdrdseeeiifkyvknthatthsaydie 832

Qy   357  --HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVG 410
             | ||||||||||||||||||||||||||||||||||:||||:|||:||     ||||:|
Db   833  vidifkieregecqrykpfkqlhnrrllwhgsrttnfagilsgglriappeapvtgymfg 892

Qy   411  KGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVI 470
             |||| |||||:|||||||||||||||||:|||||| |||||||||||||:||||||||||
Db   893  kglyfadmvsksanyyhtsq-gdplgllligevalgnmyelk-hashisrlpkgkhsvk 949

Qy   471  ARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLE 530
             :|| : :|   |:|:  :|  |  |:     || |:|::  :: :|:|||   || ||:|
Db   950  gigkttpdpsani--sldg--vdvplgtgis-sgvidtsllyneyivydiaqvnlkyllk 1004

Qy   531  I 531
             |
Db  1005  l 1005
```

```
RESULT   10
AAU29019
ID   AAU29019 standard; Protein; 1014 AA.
XX
AC   AAU29019;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human PARP-1 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis; associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
```

```
XX
PR    02-MAR-2000; 2000US-0517467.
XX
PA    (ISIS-) ISIS PHARM INC.
XX
PI    Popoff I, Cowsert LM;
XX
DR    WPI; 2001-602570/68.
DR    N-PSDB; AAS45582.
XX
PT    Antisense compound useful for treating hyperproliferative,
PT    neurological, inflammatory and autoimmune disorders and diabetes
PT    inhibits human PARP -
XX
PS    Example 13; Page 97-101; 168pp; English.
XX
CC    The invention relates to antisense oligonucleotides targeted to human
CC    PARP nucleic acid and inhibiting expression of human PARP. PARP
CC    (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC    decondensation, DNA replication, DNA repair, gene expression, malignant
CC    transformation, cellular differentiation and apoptosis. The antisense
CC    oligonucleotide inhibitors are useful for inhibiting the expression of
CC    PARP in human cells or tissues. They are also useful for treating a
CC    human with a disease associated with PARP especially hyperproliferative
CC    disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC    neurological (e.g parkinsonism, meningitis-associated intracranial
CC    complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC    arthritis) and diabetes. The present sequence is a PARP protein,
CC    the cdNA encoding which was used to design the antisense
CC    oligonucleotides.
XX
SQ    Sequence   1014 AA;

Query Match        24.9%;   Score 700;  DB 22;  Length 1014;
Best Local Similarity  33.6%;  Pred. No. 2.5e-53;
Matches 182;  Conservative 102;  Mismatches 219;  Indels 38;  Gaps 17;

Qy      1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          ::|||:::| | :| | ::   :  :: :::|::  |          : |   | |  :
Db    493 vaprgksga--alskkskgqvkeeginksekrmkltlkggaavdpdsglehsahvlekg 549

Qy     61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          |  ||  ::| | ::|:|: ||  :::|:  ||||  ||  | :|  ||:| ||||
Db    550 gkvfsatlgivdivkgtnsyYkiqlleddkenrywifrswgrvgtvigsnkleqmpsked 609

Qy    118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
          | ||| :|  |   |     |   ||:||  :| :   |   : ||::|      : |::
Db    610 aieqfmklyeektgnawhskn-ftkypkkfypleid-ygqdeeavkkl---tvnpgtksk 664

Qy    178 VVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
          :  |||   ||  | |||  ::|| | |     ::::||::| | : |  :     :::
Db    665 lpkp------vqdlikmifdvesmkkamveyeidlqkmplgklskrqiqaaysilsevqq 718

Qy    238 AMKNPTGDGQSLEELSSCCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTL 297
          |     |    :|| || ::|:| |  ||:  ||  |  ::|::: ||| ||   | |  :|
Db    719 avsggssdsgil-dlsnrfytliphdfgmkkppplnnadsvqakvemldnlldievaysl 777

Qy    298 QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR 356
          ::||      ||  |: ::| : :::||    |  ::  : | ||  : :|||| | |::|
Db    778 lrggsddssk-----dpidvnyeklktdikvvdrdseeaeliirkyvknthatthsaydle 832

Qy    357 --HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVG 410
          |  ::| |:: ||   | ||:    | ||  | ::   |::| :||:
Db    833 vidifkieregecqrykpfkqlhnrrllwhgsrttnfagilsqglriappeapvtgymfg 892

Qy    411 KGIYFASENSKSAGYVTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVI 470
          |||:| :| |: |    ||     |   | | | ||:      ::   |  |:|:
Db    893 kgiyfadnvsksanyyhtsq--gdpigllllgevalgnmyek-hashisrlpkgkhsvk 949

Qy    471 ARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLE 530
          |||  |      || :|: |  | |  |  |  :|   |:  :| :  :: || || ||
Db    950 glgkttpdpsanl--sldg--vdvplgtgis-sgvidtsllyneyivydiaqvnlkyllk 1004
```

CC   agent, and in cells that survive this treatment they inhibit genomic
CC   instability, so should reduce the likelihood of further development of
CC   both malignant cells and resistance to chemotherapy. This sequence
CC   represents the human poly(ADP-ribose) polymerase described in the
CC   invention.
XX
SQ   Sequence 1014 AA;

Query Match   24.8%; Score 699; DB 20; Length 1014;
Best Local Similarity   33.8%; Pred. No. 3.1e-53;
Matches 183; Conservative 99; Mismatches 220; Indels 40; Gaps 17;

Qy   1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          :||:|: :| ::|: | : :|:   :: | :
Db   493 vaprgksga--alskkskgqvkeeginksekrmkltlkggaavdpdsglehsahvlekg 549

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          :|:: : : | ::| :: :||| |::| || | :   : : :|: |
Db   550 gkvfsatiglvdivkgtnsyykiqlleddkenrywifrswgrvgtvigsnkleqmpsked 609

Qy   118 AKKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
          | ::|| | : || : : |: | | : | :   : : :|||| : :|
Db   610 aiehfmklyeektgnawhskn-ftkypkfypleid-yqqdeeavkkl---tvnpgtksk 664

Qy   178 VVKPCSLDPATQNLITNIFSKEMFKNAWTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
          ::| | || ||:|::|||||||:||: ::| || :|:   :::
Db   665 lpkp------vqdlikmifdvesmkkamveyeidlqkmplgklskrqiqaayslsevgq 718

Qy   238 AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTL 297
          || | ||| :||:|| ||| |::| |:||| |: :| | | :: |||||:||:
Db   719 avsggssdsqil-dlsnrfytliphdfgmkkpplinnadsvqakvemlnlldievaysl 777

Qy   298 QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPLR 356
          : : | : ||| || | : | | : :||:| :: |:|
Db   778 lrggsddssk------dpidvnyeklktdikvvdrdseeeiirkyvknthatthsaydle 832

Qy   357 --HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVG 410
            :: :|:| |: :| | | |||||| :|| :|| :||:||| | | :: |
Db   833 vidifklereqecqrykpfkqlhnrrllwhgsrtnfagilsgglriappeapvtgymfg 892

Qy   411 KGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVI 470
          ||||||:: | || | |: | :| : : | ||::| ::|:| : |: :
Db   893 kgiyfadmvsksanyyhtsq-gdpiglilgevalgnmyelk-hashisrlpkgkhsvk 949

Qy   471 ARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLE 530
          |: :|| |:| | | || | | || : |::| : | |: |::|| ||
Db   950 gigkttpdpsani--sldg--vdvpigtgis-sgvndtsllyneytvydiaqvnlkyllk 1004

Qy   531 I 531
          :
Db   1005 l 1005

RESULT 12
AAY33699
ID   AAY33699 standard; Protein; 1014 AA.
XX
AC   AAY33699;
XX
DT   18-JAN-2000 (first entry)
XX
DE   Human poly(ADP-ribose)polymerase protein.
XX
KW   Gene therapy; poly(adenosine diphosphate-ribose) polymerase; treatment;
KW   PARP; antitumor; nuclear DNA repair; proliferating cell; DNA damage;
KW   protection; genomic instability; cancer; prevention; human.
XX
OS   Homo sapiens.
XX
PN   DE19808889-A1.
XX
PD   09-SEP-1999.
XX
PF   03-MAR-1998; 98DE-1008889.
XX
PR   03-MAR-1998; 98DE-1008889.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI   Buerkle A, Meyer R;
XX
DR   WPI; 1999-509563/43.
DR   N-PSDB; AA23799.
XX
PT   New gene therapy vector expressing poly(adenosine   -
PT   diphosphate-ribose) polymerase for treating or preventing tumors
XX
PS   Disclosure; Fig 2; 12pp; German.
XX
CC   This invention describes a novel vector (A), suitable for gene therapy,
CC   which contains a DNA insert (I) that encodes a practically complete
CC   poly(adenosine diphosphate-ribose) polymerase (PARP) which has antitumor
CC   activity. PARP is a nuclear DNA repair enzyme (activated by strand
CC   breakage), that allows recovery of proliferating cells from the toxic
CC   effects of DNA damage (caused by alkylating or oxidizing agents or
CC   radiation), and protects cells against such damage or other causes
CC   of genomic instability. (A) are used for treatment of cancer and for
CC   cancer prevention in subjects at high risk (e.g. those with
CC   tumor-associated genetic defects). Tumor cells treated with (A) show an
CC   increased tendency to die when treated with radiation or chemotherapeutic

RESULT 13
AAB60693
ID   AAB60693 standard; Protein; 521 AA.
XX
AC   AAB60693;
XX
DT   11-MAY-2001 (first entry)
XX
DE   Human poly(ADP-ribose) synthetase sbhPARS2.
XX
KW   Human; poly(ADP-ribose) synthetase; sbhPARS2; vaccine; drug screening;
KW   ischaemic disorder; cerebral ischaemia; cardiac ischaemia;
KW   myocardial infarction; stroke; inflammation; autoimmune disease;
KW   diabetes; multiple sclerosis; neurodegenerative disease;
KW   Parkinson's disease; Alzheimer's disease; chromosome localisation.
XX
OS   Homo sapiens.
XX
PN   WO20011264S-A1.

```
Qy   352 CPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVAAILTSGLRIMPH----SGG 407
             |:  :|| :|||  ||||||| |  || |||||||  ::  || :|| : || :
Db   336 tmtldlfevekdgekeafr--edlhnrmllwhgsrmsnwvgilshglriappeapitgy 393

Qy   408 RVGKGIYFASENSKSAGYVTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFD 467
             ||||||| :||||   |         |  | ||||||| ||||    :: :  :  :
Db   394 mfgkgiyfadmssksanycfasrl--kntglillsevalgqcnelleanpkaegllqgkh 451

Qy   468 SVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSF--SQSEYLIYKESQCRL 525
             | ||: | :  :||||   || |  |||| : |||| ::  :||: ::|| :|: |:
Db   452 stkglgkmapssahfv--tlngs--tvplgpasdtgilnpdgytlhyneyivynpnqvrm 507

Qy   526 RYLLEI 531
             |||::
Db   508 ryllkv 513

RESULT 14
AAY51174
ID   AAY51174 standard; Protein; 570 AA.
XX
AC   AAY51174;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human brain PARP2 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage.
XX
OS   Homo sapiens.
XX
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44287.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 52-54; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP2 protein used in the method of the invention.
XX
SQ   Sequence    570 AA;
```

```
XX
PD   22-FEB-2001.
XX
PF   10-AUG-2000; 2000WO-US21775.
XX
PR   12-AUG-1999;   99US-0373441.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Barone F, Field J, Kabnick KS, Li X, McQueney MS, Zhu Y;
XX
DR   WPI; 2001-211196/21.
DR   N-PSDB; AAF59996.
XX
PT   New human polypeptide of the polyADPribose synthetase family for use in
PT   diagnostic assays and for screening modulators used for preventing and
PT   treating inflammation, autoimmune disease and Alzheimers disease  -
XX
PS   Claim 1; Page 21; 30pp; English.
XX
CC   This sequence represents a novel human poly(ADP-ribose)
CC   synthetase, sbhPARS2. The invention also relates to fragments,
CC   variants and sequences with at least 95% identity to the sbhPARS2
CC   protein or nucleotide sequence; expression systems and host cells
CC   comprising an sbhPARS2 nucleic acid sequence; the recombinant expression
CC   of sbhPARS2; and an antibody specific for sbhPARS2. sbhPARS2 proteins
CC   and nucleotides are useful as vaccines for inducing an immunological
CC   response in a mammal. The sbhPARS2 protein is useful for identifying
CC   compounds which inhibit or stimulate its activity or expression level.
CC   Such agonists and antagonists of sbhPARS2 are useful for treating human
CC   diseases including ischaemia and ischaemic tissue injury (e.g., cerebral
CC   and cardiac ischaemia, myocardial infarction, stroke), inflammation,
CC   autoimmune disease (e.g. diabetes, multiple sclerosis) and
CC   neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's
CC   disease). sbhPARS2 nucleic acids are useful as diagnostic reagents for
CC   detecting mutations in the associated gene; as hybridisation probes
CC   to isolate full-length sbhPARS2 cDNAs and sbhPARS2 genomic clones; and
CC   for chromosome localisation studies. The sbhPARS2 protein is also useful
CC   as an immunogen to produce antibodies for therapeutic use. sbhPARS2
CC   proteins, nucleotides and antibodies are also useful in screening methods
CC   for detecting the effect of added compounds on the production of mRNA and
CC   protein in cells.
XX
SQ   Sequence    521 AA;

Query Match             24.8%;  Score 698;  DB 22;  Length 521;
Best Local Similarity   33.9%;  Pred. No. 1.3e-53;
Matches  185;  Conservative  92;  Mismatches  205;  Indels  64;  Gaps  18;

Qy    13 GSKKQRQGTEEEDSFRSTAEALRA-APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTL 68
             | | || :: :     |  |: ||      |        ||  |:  ::  ||
Db     5 ggkankdrtedkqdesvkallkgkap------vdpectakvgkahvycegndvydvml 57

Qy    69 NQTNIGNNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFK 123
             |||| ||:|:|||||:: ::  || | |||||||||||::|| |  |:|:|:
Db    58 nqtnlqfnnnkyyilqlleddagrnfsvmrwgrvgkmqghsl--vacsgnlnkakeifq 115

Qy   124 KFWEKTKNKWEERDRFVAQPNKYTLIEV-------QGEAESQEAVVKALSPQVDSGPV 175
             |||| |||| ||::|| ||::|| |:::|      :  ::  :||| |::||
Db   116 kkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpe------ 169

Qy   176 RTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAL 235
             |  : ||  | || :|  :| |||| |  | |||||:|||| |:: : |: ::|:|:
Db   170 ------sqdlrvqeliklicnvqameenmumemkyntkkaplgkltvaqikadyqslkki 223

Qy   236 EEAMKNPTGD-GQSLEELSSCFYIVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELA 294
             :: ::    :|| :||||| |||: |  || |:|:| |:|:|| |:|:| :||:|||:
Db   224 edcir--agghgralmeacnefytriphdfglirtpplirtqkelsekiqllealgdiela 281

Qy   295 QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYR 351
             :::: |   : ::    | || :  | || ||  ||:|:|:||| ::  :  |  |||
Db   282 ikl-----vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdy 335
```

Query Match   24.8%; Score 698; DB 21; Length 570;
Best Local Similarity 33.9%; Pred. No. 1.5e-53;
Matches 185; Conservative 92; Mismatches 205; Indels 64; Gaps 18;

Qy   13 GSKKQRQGTEEEDSFRSTAEALRA-APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTL 68
          | |:: || :: | ||| | : ::|| :||:|:
Db   54 ggkankdrtedkqdesvkallikgkap------vdpectakvgkahvycegndvydvml 106

Qy   69 NQTNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFK 123
          |||| ||||::|:|||||::: - | |||||||:|: | ::| :: |:|:
Db   107 nqtnlqfnnnkyyliqlleddaqrnfsvmwrwgrvgkmgqhsl--vacsgnlnkakeifq 164

Qy   124 KKFWEKTKNKWEERDRFVAQPNKYTLIEV-------QGEAESQEAVVKALSPQVDSGPV 175
          ||| |:|||||||||||||: || ::::   ::| | :|
Db   165 kkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpe------ 218

Qy   176 RTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAL 235
          | || || |||| | |||||| ||||||||||| ||||||||||| ||| ||
Db   219 ------sqldlrvgeliklicnvqameemmemkyntkkaplgkltvaqikagyqslkki 272

Qy   236 EEAMKNPTGD-GQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELA 294
          | ::: | || :| | |||||| ||||||||||: |: |:|||| ||| |:|
Db   273 edcir--agghgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieia 330

Qy   295 QTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYR 351
          | || |||| | || ::| ::| : : || ::|
Db   331 1kl------vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdy 384

Qy   352 CPNLRHVWVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGG 407
          ||||| | :: | :| | || :| ||| |||| :: : || ||:|| :|
Db   385 tmtlldlfevekdgekeafr--edlhnrmllwhgsrmsnwvglishglriappeapitgy 442

Qy   408 RVGKGIYFASENSKSAGIYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFD 467
          | |||:|||: |:|: || : || | |:: | : : | |
Db   443 mfgkgiyfadmssksanycfasrl--kntglllsevalggcnelleanpkaegllqgkh 500

Qy   468 SVIARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSF--SQSEYLIYKESQCRL 525
          | | :: :||| |: ::| :|: :: || |
Db   501 stkglgkmapssahfv--tlngs-tvplgpasdtglinpdgytlnyneyivynpnqvrm 556

Qy   526 RYLLEI 531
          ||:||:
Db   557 ryllkv 562

RESULT 15
AAB11480
ID   AAB11480 standard; Protein; 570 AA.
XX
AC   AAB11480;
XX
DT   02-MAR-2001 (first entry)
XX
DE   Human brain poly-ADP-ribose-polymerase protein.
XX
KW   Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
KW   cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
KW   anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
KW   antiinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
KW   '2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
KW   neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
KW   Parkinson's disease; ischemic damage; microinfarction; sepsis;
KW   diabetes mellitus.
XX
OS   Homo sapiens.
XX
PN   DE19921567-A1.
XX
PD   16-NOV-2000.
XX
PF   11-MAY-1999; 99DE-1021567..
XX

PR   11-MAY-1999; 99DE-1021567.
XX
PA   (BADI) BASF AG.
XX
PI   Lubisch W, Sadowski J, Kock M, Hoeger T;
XX
DR   WPI; 2001-032983/05.
DR   N-PSDB; AAC82090.
XX
PT   Drugs for inhibiting PARP or especially homologous enzymes comprising
PT   4-substituted phthalazinone derivatives, useful e.g. for treating
PT   neurodegenerative disease, ischemic damage, tumors or diabetes -
XX
PS   Example A; Page 12-13; 14pp; German.
XX
CC   This invention describes novel 4-substituted 2H-phthalazin-1-one
CC   derivatives (I) which are used for the treatment or prophylaxis of
CC   diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;
CC   EC 2.4.2.30) activity. The products of the invention have nootropic,
CC   neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,
CC   cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,
CC   immunosuppressive, antiinflammatory, antirheumatic, antiarthritic,
CC   antidiabetic. (I) are especially used for treating or preventing
CC   neurodegenerative disease or neuronal damage (specifically associated
CC   with ischemia, trauma or massive bleeding, especially apoplexy or
CC   spinal-cranial trauma, or Alzheimer's disease, Huntington's disease or
CC   Parkinson's disease), treating or preventing ischemic damage
CC   (specifically renal damage after renal ischemia or during and after
CC   kidney transplantation or heart damage after cardiac ischemia), treating
CC   epilepsy, specifically generalized epileptic attacks (e.g. petit mal and
CC   tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lope
CC   and complex partial attacks), treating microinfarction (e.g. during and
CC   after heart valve replacement, aneurysm resectioning and heart
CC   transplantation), revascularization of critically constricted coronary
CC   arteries (e.g. after PCTA or by-pass operations) or peripheral arteries
CC   (e.g. leg arteries), treating acute myocardial ischemia and damage during
CC   or after its mechanical or drug-induced lysis and treating tumors and
CC   their metastasis, sepsis and septic shock, inflammatory and rheumatic
CC   disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)
CC   inhibit PARP (i.e. the known form designated PARP1), they especially
CC   selectively and strongly inhibit PARP homologs, specifically the homolog
CC   PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.
CC   with K_i values of 1-20 nM) and high selectivity for PARP2 relative to
CC   PARP1 (generally by a factor of more than 5).
XX
SQ   Sequence   570 AA;

Query Match   24.8%; Score 698; DB 22; Length 570;
Best Local Similarity 33.9%; Pred. No. 1.5e-53;
Matches 185; Conservative 92; Mismatches 205; Indels 64; Gaps 18;

Qy   13 GSKKQRQGTEEEDSFRSTAEALRA-APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTL 68
          | |:: || :: | ||| | : ::|| :||:|:
Db   54 ggkankdrtedkqdesvkallikgkap------vdpectakvgkahvycegndvydvml 106

Qy   69 NQTNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFK 123
          |||| ||||::|:|||||::: | |||||||:|: | ::| :: |:|:
Db   107 nqtnlqfnnnkyyliqlleddaqrnfsvmwrwgrvgkmgqhsl--vacsgnlnkakeifq 164

Qy   124 KKFWEKTKNKWEERDRFVAQPNKYTLIEV-------QGEAESQEAVVKALSPQVDSGPV 175
          ||| |:|||||||||||||: || ::::   ::| | :|
Db   165 kkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpe------ 218

Qy   176 RTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAL 235
          | || || |||| | |||||| ||||||||||| ||||||||||| ||| ||
Db   219 ------sqldlrvgeliklicnvqameemmemkyntkkaplgkltvaqikagyqslkki 272

Qy   236 EEAMKNPTGD-GQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELA 294
          | ::: | || :| | |||||| ||||||||||: |: |:|||| ||| |:|
Db   273 edcir--agghgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieia 330

Qy   295 QTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYR 351

```
Db   331 ikl------vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdy 384
             :.:::   ||||.|:.|.||.|| |:|.|  ||:.|  ||: |   .:|

Qy   352 CPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGG 407
             ...:|..::||.|: :|| ||:|| |||:.:  ||.||||   | :)

Db   385 tmtldlfevekdgekeafr--edlhnrmllwhgsrmsnwvgilshglriappeapltgy 442

Qy   408 RVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFD 467
             |||||||| :|||| |  : |:||||: :.: ..:  .:  : :  |

Db   443 mfgkgiyfadmssksanycfasrl--kntglllsevalgqcnelleanpkaegllgkh 500

Qy   468 SVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSF--SQSEYLIYKESQCRL 525
             |-:.|-|-|-.:.| .|:| -||.||  .:.||:.|.| |:.|:|:

Db   501 stkglgkmapssahfv--tlngs--tvplgpasdtgilnpdgytlnyneyivynpnqvrm 556

Qy   526 RYLLEI 531
             |||:::

Db   557 ryllkv 562


Search completed: August 29, 2002, 07:57:19
Job time: 296 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 29, 2002, 07:59:33 ; Search time 69.02 Seconds
                                         (without alignments)
                                         742.040 Million cell updates/sec

Title:          US-09-701-586b-8
Perfect score:  2813
Sequence:       1 MAPKRKASVQTEGSKKQRQG.........EVLIYKESQCRLRYLLEIHL 533

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Database :      PIR_71:*
                1: pir1:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

Searched:       283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters:    283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1975	70.2	459	2	T08713	NAD+ ADP-ribosyltr
2	717	25.5	653	2	T03656	probable NAD+ ADP-
3	705	25.1	1014	1	A29725	NAD+ ADP-ribosyltr
4	702	25.0	1016	1	JS0428	NAD+ ADP-ribosyltr
5	686.5	24.4	998	2	S31735	NAD+ ADP-ribosyltr
6	686.5	24.4	1011	1	JH0581	NAD+ ADP-ribosyltr
7	683.5	24.3	635	2	T01311	NAD+ ADP-ribosyltr
8	683.5	24.3	1013	1	S04200	NAD+ ADP-ribosyltr
9	682.5	24.3	500	2	S26057	NAD+ ADP-ribosyltr
10	645.5	22.9	996	1	S42208	NAD+ ADP-ribosyltr
11	621.5	22.1	983	2	T51353	NAD+ ADP-ribosyltr
12	621.5	22.1	1009	2	C84719	probable poly (ADP
13	619	22.0	994	1	A47474	NAD+ ADP-ribosyltr
14	584.5	20.8	969	2	T03657	NAD+ ADP-ribosyltr
15	475.5	16.9	727	2	T18600	hypothetical prote
16	396.5	14.1	538	2	T20414	hypothetical prote
17	193	6.9	135	2	PN0494	NAD+ ADP-ribosyltr
18	180	6.4	2004	2	D88948	protein ZK1005.1 [
19	125.5	4.5	5175	2	T20992	hemicentin precurs
20	125.5	4.5	5198	2	T43290	hemicentin precurs
21	120.5	4.3	181	2	T03058	NAD+ ADP-ribosyltr
22	119.5	4.2	954	2	E86174	protein F19P19.26
23	118.5	4.2	1092	2	T33717	carbamoyl-phosphat
24	114.5	4.1	1203	2	D86625	subunit A of ATP-d
25	112.5	4.0	935	2	E96806	hypothetical prote
26	111.5	4.0	359	2	T12540	hypothetical prote
27	110	3.9	1223	2	T17345	hypothetical prote
28	108	3.8	949	2	D97781	hypothetical prote
29	106.5	3.8	638	2	C83503	exotoxin A precurs
30	106.5	3.8	1031	1	A38713	kinesin heavy chai
31	106.5	3.8	1435	2	S69632	regulatory protein
32	106	3.8	769	2	E96613	hypothetical prote
33	105.5	3.8	4162	2	T42633	connectin/titin -
34	105	3.7	511	2	S38905	hypothetical prote
35	105	3.7	3187	2	JC5837	364K Golgi complex
36	104.5	3.7	963	1	A41919	kinesin heavy chai
37	104.5	3.7	1173	2	T25539	hypothetical prote
38	104	3.7	2484	2	T26216	hypothetical prote
39	104	3.7	2607	2	T26215	hypothetical prote
40	103.5	3.7	638	2	A30347	exotoxin A precurs
41	102.5	3.6	1280	2	T00365	hypothetical prote
42	102	3.6	881	2	I84737	kinesin heavy chai
43	101.5	3.6	1244	2	A05218	hypothetical prote
44	101	3.6	783	2	T18421	hypothetical prote
45	101	3.6	3075	2	S14458	laminin alpha-1 ch

ALIGNMENTS

```
RESULT    1
T08713
NAD+ ADP-ribosyltransferase homolog DKFZp566G0224.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Nov-2000
C;Accession: T08713
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16472
A;Accession: T08713
A;Molecule type: mRNA
A;Residues: 1-459 <ANS>
A;Cross-references: EMBL:AL050034
A;Experimental source: fetal kidney; clone DKFZp566G0224
C;Genetics:
A;Note: DKFZp566G0224.1
C;Superfamily: NAD+ ADP-ribosyltransferase


Query Match             70.2%;  Score 1975;  DB 2;  Length 459;
Best Local Similarity   81.5%;  Pred. No. 2.3e-135;
Matches  379;  Conservative  32;  Mismatches  44;  Indels  10;  Gaps  4;

Qy   73  IGNNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLEDAKKDFKKKFWEKTK 131
             ||||||||||||:||||| :||||||||||||||||:||||||||||:|||||
Db    1  IENNNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFKKKFREKTK  60

Qy  132  NKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTV---VKPCSLDPAT 188
             :|| |||| :: |||||||| :|||||  || ||||||  :||||||   :|||||||||
Db   61  NNWAERDHFVSHPGKYTLIEVQEADEAQEAVVK-----VDRGPVRTVTKRVQPCSLDPAT 115

Qy  189  QNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQS 248
             |:||||||||||||||||||||||||||||:||||||||||||||||||:|| |||||
Db  116  QKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEALKGPTDGGQS 175

Qy  249  LEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKV 308
             |||||: |||||||||||||||||||:|||||||||||||||||||:|| :|: |
Db  176  LEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQAV-SEQEKTV 234

Qy  309  EEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWKVNREGEGD 368
             ||||||||||||||||:|||||||||| ||:|||:|:|:::|||:|||||||||:|||
Db  235  EEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVNQEGEED 294

Qy  369  RFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGYVTT 428
             |||||||||||||||:||:|:||||||||||||||||||||:|||||||||||||||||
Db  295  RFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGRVGEGIYFASENSKSAGYVIG 354

Qy  429  MHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIELELD 488
             | ||| ||||||||||| ||||:|| |||::|||||||||||:||| ||| || ||||||
Db  355  MKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELD 414

Qy  489  GQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 533
```

```
Db    415 GQQVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 459
          || |||||  || ||  |||||||||::|||||||||||:||

RESULT    2
T03656
Probable NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Oct-1999
C;Accession: T03656
R;Babylchuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Monta
submitted to the EMBL Data Library, November 1997
A;Description: Higher plants possess two poly(ADP-ribose) polymerases.
A;Reference number: Z14991
A;Accession: T03656
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-653 <BAB>
A;Cross-references: EMBL:AJ222588; NID:e1264090; PIDN:CAA10888.1; PID:e1264091
C;Genetics:
A;Gene: PARP
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase


Query Match   25.5%;   Score 717;  DB 2;    Length 653;
Best Local Similarity  34.9%;   Pred. No. 5.8e-44;
Matches  190;  Conservative  91;  Mismatches  205;  Indels  58;  Gaps  17;

Qy    13 GSKKQRQGTEEEDSFRS----IRVDPSCPFSRNPGIQVHED-Y 64
         |||||||||||||||  :.:      |||  |||||  ||  :
Db   138 GAKEVIKGGDEEVEVKKEKMVTATKKGAAVLDQHIPDHIKVNYHV------LQVGDEIY 190

Qy    65 DCTLNQTNIGNNNNKFYIIQLLEE--GSRPFCWNRWGRVGEVGQSKMNHFTCLED-AKKD 121
         ||||||||||||||||||||||||    | |||||||||||  | ||||||     | |
Db   191 DATLNQTNVGDNNNKFYIIQVLESDAGGSFMVYNRWGRVGRGQDKLHGPSPTRDQAIYE 250

Qy   122 FKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKALSPQVDSGPVRTVVK 180
         | | |||  | | || :| | :||| ||| || ||   | | ||||||||||||||||||
Db   251 FEGKPHNKTNNHWSDRKNFKCYAKKYTWLEMDYGETEK--------EIEKGSITDQIK 300

Qy   181 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 240
         |||||||||||||||||||||||||||||||  || :||||||||||||||||  ||||:
Db   301 ETKLETRIAQFISLICNISMMKQRMVEIGYNAEKLPLGKLRKATILKGYHVLKRISDVIS 360

Qy   241 NPTGDGQSLEELSSCFYTVIPHNFG--RSRPPTNSPDVLQAKKDMLLVLADIELAQTLQ 298
         |||||||||||||||||| :||||||  |:| |  |  |::|||  :::  |  |||| |
Db   361 K--ADRRHLEQLTGEFYTVIPHDFGFRKMREFIIDTPQKLKAKLEMVEALGEIBIATKLL 418

Qy   299 AAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT--GNSYRCPNL 355
         | |||| || ||  |  | ||   |||  :||||:|  ||| |  ||:  || ||||||
Db   419 ------EDDSSDQDDPLYARYKQLHCDFTPLEADSDEYSMIKSYLRNTHGKTHSGYTVDI 472

Qy   356 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK 411
         || |  || :| |  | ||| | || |||  |  | | ||||||| :|     ||||||||
Db   473 VQIFKVSRHGETERFQKFASTRNRMLLWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGK 532

Qy   412 GIYFASENSKSAGVTTMHCGGHQV--GYMFIGEVALGKEHHITIDDPSLKSPPPGFDS 468
         ||||||||||    :||| :|||::     ||| ||||| :|||| ||||||| |||||
Db   533 GVYFADMFSKSANY----CYASEACRSGVLLLCEVALGDMNELLNADYDANNLPKGKLR 587

Qy   469 VIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYL 528
         ||  |     :|      |    ||  | |  | | ||| ||||||||| |:| ||||| |
Db   588 SKGVGQTAPNM--VESKVADDGVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMRYV 643

Qy   529 LEIH 532
         |:||
Db   644 LHVN 647


RESULT    3
A29725
NAD+ ADP-ribosyltransferase (EC 2.4.2.30), nuclear - human
N;Alternate names: poly (ADP-ribose) polymerase; poly (ADP-ribose) synthetase; poly(ADP)
```

```
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29725; A28498; A39976; A26901; I38096; B33321; A33321; A35635; A61559;
R;Uchida, K.; Morita, T.; Sato, T.; Ogura, T.; Yamashita, R.; Noguchi, S.; Suzuki, H.
Biochem. Biophys. Res. Commun. 148, 617-622, 1987
A;Title: Nucleotide sequence of a full-length cDNA for human fibroblast poly(ADP-ribo
A;Reference number: A29725; MUID:88076933
A;Accession: A29725
A;Molecule type: mRNA
A;Residues: 1-69,'Q',71-1014 <UCH>
A;Cross-references: GB:M18112; NID:g190166; PIDN:AAA60137.1; PID:g190167
R;Kurosaki, T.; Ushiro, H.; Mitsuuchi, Y.; Suzuki, S.; Matsuda, M.; Matsuda, A.; Katu
J. Biol. Chem. 262, 15990-15997, 1987
A;Title: Primary structure of human poly (ADP-ribose) synthetase as deduced from cDNA
A;Reference number: A28498; MUID:88058958
A;Accession: A28498
A;Molecule type: mRNA
A;Residues: 1-16,'E',18-211,'K',213-236,'R',238-366,'H',369-1014 <KUR>
A;Cross-references: GB:J03473
R;Cherney, B.W.; McBride, O.W.; Chen, D.; Alkhatib, H.; Bhatia, K.; Hensley, P.; Smul
Proc. Natl. Acad. Sci. U.S.A. 84, 8370-8374, 1987
A;Title: cDNA sequence, protein structure, and chromosomal location of the human gene
A;Reference number: A39976; MUID:88068596
A;Accession: A39976
A;Molecule type: mRNA
A;Residues: 1-49,'D',51-612,'Q',614-907,'Y',909-939,'R',941-979,'I',981-1014 <CHE>
A;Cross-references: GB:J03030
A;Note: the authors translated the codon ATA for residue 980 as Asn
R;Suzuki, H.; Uchida, K.; Shima, H.; Sato, T.; Okamoto, T.; Kimura, T.; Miwa, M.
Biochem. Biophys. Res. Commun. 146, 403-409, 1987
A;Title: Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and expressi
A;Reference number: A26901; MUID:87298455
A;Accession: A26901
A;Molecule type: mRNA
A;Residues: 441-610,'N',612-880;921-1014 <SUZ>
A;Note: the sequence figure has an omission of forty residues
R;Ogura, T.; Nyunoya, H.; Takahashi-Masutani, M.; Miwa, M.; Sugimura, T.; Esumi, H.
Biochem. Biophys. Res. Commun. 167, 701-710, 1990
A;Title: Characterization of a putative promoter region of the human poly(ADP-ribose)
A;Reference number: I38096; MUID:90211250
A;Accession: I38096
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-40 <RES>
A;Cross-references: EMBL:X16674; NID:g510112; PIDN:CAA34663.1; PID:g1017423
R;Auer, B.; Nagl, U.; Herzog, H.; Schneider, R.; Schweiger, M.
DNA 8, 575-580, 1989
A;Title: Human nuclear NAD(+) ADP-ribosyltransferase(polymerizing): organization of t
A;Reference number: A33321; MUID:90091744
A;Accession: B33321
A;Molecule type: DNA
A;Residues: 38-43;93-98;132-137;204-209;237-242;276-281;335-340;384-389;431-436;512-5
A;Cross-references: GB:M29544; GB:M22953
A;Note: the authors translated the codon GTG for residue 54 as Glu
A;Note: these fragments represent intron-exon boundaries
A;Accession: A33321
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 16-66;96;121-159,'D',161-167 <AU2>
A;Note: these fragments represent a zinc finger-containing DNA-binding region
R;Gradwohl, G.; Menissier de Murcia, J.; Molinete, M.; Simonin, F.; Koken, M.; Hoeijm
Proc. Natl. Acad. Sci. U.S.A. 87, 2990-2994, 1990
A;Title: The second zinc-finger domain of poly(ADP-ribose) polymerase determines spec
A;Reference number: A35635; MUID:90222155
A;Accession: A35635
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 12-26,'T',28-66;116-166 <GRA>
R;Schneider, R.; Auer, B.; Kuehne, C.; Herzog, H.; Klocker, H.; Burtscher, H.J.; Hirs
Eur. J. Cell Biol. 44, 302-307, 1987
A;Title: Isolation of a cDNA clone for human NAD (+): protein ADP-ribosyltransferase.
A;Reference number: A61559; MUID:88082900
A;Accession: A61559
```

A;Molecule type:mRNA
A;Residues: 381-420;682-710 <SCH>
R;Yokoyama, Y.; Kawamoto, T.; Mitsuuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Terashi
Eur. J. Biochem. 194, 521-526, 1990
A;Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.
A;Reference number: S14010; MUID:91099327
A;Accession: S14010
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-95 <YOK>
A;Cross-references: EMBL:X56140; NID:g35286; PIDN:CAA39606.1; PID:g825702
C;Comment: This protein can ADP-ribosylate itself as well as other proteins.
C;Genetics:
A;Gene: GDB:ADPRT; PPOL
A;Cross-references: GDB;119508; OMIM:173870
A;Map position: 1q41-1q42
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransfe

```
Query Match            25.1%;   Score 705;  DB 1;  Length 1014;
Best Local Similarity  33.8%;   Pred. No. 8.1e-43;
Matches 183;  Conservative 101;  Mismatches 218;  Indels 40;  Gaps 17;

Qy    1 MAPKRRASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQYV 60
          :||  |  :| :  :| : :| : :
Db  493 VAPRGKSGA--ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 549

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
Db  550 GKVFSATLGLVDIVKGTNSYKLQLLEDDKENRTWIFRSWRGVGVTVIGSNKLEQMPSKED 609

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
Db  610 AIEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL--TVNPGTKSK 664

Qy  178 VVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALALEE 237
Db  665 LPKP-----VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQ 718

Qy  238 AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTL 297
Db  719 AVSQSQSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSL 777

Qy  298 QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCP 353
Db  778 LRGGSDDSSK-----DPIDVNYEKLKTDIKVVDRSEEAEIIRKYVKNTHATTHNAYDL- 831

Qy  354 NLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRV 409
Db  832 EVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMF 891

Qy  410 GKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSV 469
Db  892 GKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEAALGNMYELK-HASHISKLPKGKHSV 948

Qy  470 IARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLL 529
Db  949 KGLGKTTPDPSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEYIVYDIAQVNLKYLL 1003

Qy  530 EI 531
Db 1004 KL 1005
```

submitted to JIPID, February 1990
A;Reference number: JS0428
A;Accession: JS0428
A;Molecule type: mRNA
A;Residues: 1-1016 <SAI>
A;Experimental source: thymus
R;Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; T
Eur. J. Biochem. 171, 571-575, 1988
A;Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int
A;Reference number: S00328; MUID:88151954
A;Accession: S00328
A;Molecule type: mRNA
A;Residues: 648-714;838-904 <TAN>
A;Cross-references: EMBL:X06986
A;Accession: A30458
A;Molecule type: protein
A;Residues: 658-685;689-696;893-901 <TA2>
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-loo
F;21-51/Region: zinc finger
F;128-165/Region: zinc finger
F;200-220/Region: helix-turn-helix motif
F;224-231/Region: nuclear location signal
F;250-270/Region: helix-turn-helix motif
F;494-501/Region: nucleotide-binding motif A (P-loop)
F;890-903/Region: nucleotide binding #status predicted

```
Query Match            25.0%;   Score 702;  DB 1;  Length 1016;
Best Local Similarity  34.2%;   Pred. No. 1.3e-42;
Matches 187;  Conservative 95;  Mismatches 215;  Indels 20;  Gaps 20:

Qy    1 MAPKRRASVQTEGSKKQRQGTEEEDSFRSTAE--ALRAAPADNRVIRVDPSCPFSRNPG 57
Db  495 VGPKGKSGAAP--SKKSKGPVKEEGTNKSEKRMKLTLKGGAA------VDPDSGLEHNAH 546

Qy   58 I--QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHF 112
Db  547 VLEKGKGVFSATLGLVDIVKGTNSYKLQLLEDDKESRYWIFRSWGRVGVTVIGSNKLEQM 606

Qy  113 TCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDS 172
Db  607 PSKEDAIEHFMKLYEEKTGNAWHSKN-FTKHPKKFYPLEID-YGQDEEAVKKL--TVNP 661

Qy  173 GPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL 232
Db  662 GTKSKLPKP------VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSIL 715

Qy  233 EALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIE 292
Db  716 SEVQQALSQGSGSDSHIL-DLSNRFYTLIPHDFGMKKPPLLNNANSVQAKVEMLDNLLDIE 774

Qy  293 LAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGN 348
Db  775 VAYSLLRGGSDDSSK------DPIDVNYEKLKTDIKVVDKDSEEAEIIRKYVKNTHATTHN 829

Qy  349 SYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---- 404
Db  830 AYDL-EVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPV 888

Qy  405 SGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPP 464
Db  889 TGYMFGKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEAALGNMYELK-HARHISKLPK 945

Qy  465 GFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCR 524
Db  946 GKHSVKGLGKTTPDPSASI--TVDG--VEVPLGTGIS-SGVNDTCLLYNEYIVYDIAQVH 1000

Qy  525 LRYLLEI 531
Db 1001 LKYLLKL 1007
```

RESULT   4
JS0428
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine
N;Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate rib
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: JS0428; S00328; A30458
R;Saito, I.

RESULT 5
S31735
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - African clawed frog (fragment)
N;Alternate names: poly ADP-ribose polymerase
C;Species: Xenopus laevis (African clawed frog)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S31735; PN0495
R;Saulier-le Drean, B.M.
submitted to the EMBL Data Library, May 1992
A;Reference number: S31735
A;Molecule type: mRNA
A;Accession: S31735
A;Residues: 1-998 <SAU>
A;cross-references: EMBL:Z12139; NID:g64967; PIDN:CAA78126.1; PID:g1334661
R;Ozawa, Y.; Uchida, K.; Ami, Y.; Kushida, S.; Okada, N.; Miwa, M.
Biochem. Biophys. Res. Commun. 193, 119-125, 1993
A;Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)polymerase
A;Reference number: PN0494; MUID:93277538
A;Accession: PN0495
A;Molecule type: mRNA
A;Residues: 742-745,'E', 747-876 <O2A>
C;Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and differe
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pentosyl

Query Match   24.4%;   Score 686.5;   DB 2;   Length 998;
Best Local Similarity   34.1%;   Pred. No. 1.7e-41;
Matches 182;   Conservative 95;   Mismatches 212;   Indels 45;   Gaps 19;

Qy   10 QTEGSKKQRQGT-EEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHEDYDCTL   68
          :| || :|| :|:| : || : :|| ::   : |||||||||||| ||
Db   488 KSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDS---ELEDSCHVLETGG----KIFSATL   540

Qy   69 NQTNIGNNNNKFYIIQLLE--EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKKDFKKK   125
          || :|:||||||| :: :| | : : ||| |||||| ||| |||||| ||| | :
Db   541 GLVDITRGTNSYYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKEDAIEHFLNL   600

Qy   126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLD   185
          : || :||::|| :| : :: ||||| | | ::::| | | :|| : | ||
Db   601 YQDKTGNAWHS-PNFTKYPKKFYPLEI--DYGQEEDVVKKLS--VGAGTKSKLAKP----   651

Qy   186 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQOIARGFEALEALEEAMKNPTGD   245
          | | ||| :|||:||||||||||||||||:|: ||||||| | : | :|| |
Db   652 --VQELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAVSESLSE   709

Qy   246 GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE   305
          |:|||||||||||||||:||| :||||||| |:: : ||| ||| :| :||
Db   710 ARLL-DLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDG   768

Qy   306 EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLK----QTGNSYRCPNLRHVWKV   361
          || | :| |: : : || || :| :|| | || : | :| ||| |
Db   769 EK-----DPIDVKYEKIKTDIKVVAKDSEESRIICDVVKNTHADTHNAYDLEVL-EIFKI   822

Qy   362 NREGEGDRRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFAS   417
          |||||| ::|| || : :||:|||:||||||||||||| || ||| | ||||||
Db   823 DREGEYQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFAD   882

Qy   418 ENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEP   477
          ||||:||||||||||||||||:||| :|:|||:| ||||||||| :| || : |
Db   883 MVSKSANYCHAM--PGSPIGLILLGEVALGNMHELKAASQITKL-PKGKHSVKGLGRTAP   939

Qy   478 DPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI   531
          || | | | |: || | :: || ||:: | | ||||||||||::
Db   940 DPSATV--QLDG--VDVPLGKGYSA-NLSDTSLLYNEYIVYDIAQVNLKYLLKL   988

RESULT 6
JH0581
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - chicken
N;Alternate names: poly(ADP-ribose) synthase
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: JH0581
R;Ittel, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.
Gene 102, 157-164, 1991
A;Title: Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence an
A;Reference number: JH0581; MUID:91340148
A;Accession: JH0581
A;Molecule type: mRNA
A;Residues: 1-1011 <ITT>
A;Cross-references: EMBL:X52690; NID:g63742; PIDN:CAA36917.1; PID:g63743
C;Comment: This protein is a chromatin-bound enzyme.
C;Comment: This enzyme catalyzes DNA-dependent post-translational modifications of va
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc

Query Match   24.4%;   Score 686.5;   DB 1;   Length 1011;
Best Local Similarity   33.8%;   Pred. No. 1.8e-41;
Matches 184;   Conservative 97;   Mismatches 207;   Indels 57;   Gaps 20;

Qy   5 RKASVQTEGSKKQRQG-TEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED   63
          :| :|| |||::|: | || : :|| ::   : |||||||||
Db   497 KPANMKSAGKVKEEQGPSKSEKKMKLTVKGGAAVDPDSGL--EDSAHVFEKGGKI----   549

Qy   64 YDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKK   120
          || ::|: |:||||||| :: :|| | : : ||| |||||| ||| ||||| |||
Db   550 FSATIGLVDIVKGTNSYYKLQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAVE   609

Qy   121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVK   180
          :: || || :||::|| :| : :: ||||| | | ::::| | | :|| : |
Db   610 HFLNLYEEKTGNSWHSKN-FTKYPKKFYPLEID-YGQDEEAVRKL---TVSAGTKSKLAK   664

Qy   181 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK   240
          | | | | ||| :|||:||||||||||||||||||: ||||||| | : | :|| |
Db   665 P------IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILNEVQQAVS   718

Qy   241 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAA   300
          | ::| |:|||||||||||||||:||| :||| ||| |:: : ||| ||| :| :||
Db   719 DGGSESQIL-DLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRG   777

Qy   301 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLR   356
          ||: || | :| |: : : || || || :| :|| | ||:   :| ||| |:
Db   778 GNEDGDK------DPIDINYEKLRTDIKVVDKDSEEAKIIKQYVKNTHAATHNAYDL-KVV   831

Qy   357 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH-----SGGRVGKG   412
          ||| |||||| ::|| || : :||:|||:||||||||||||| || ||| | ||
Db   832 EIFRIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKG   891

Qy   413 IYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEH-------HITIDDPSLKSPPPGF   466
          |||| ||||:|||||||||||||||:||| :|:||| | ||||||||| :|
Db   892 IYFRDMVSKSANYCHTSQ--ADPIGLILLGEVALGNMYELKNASHIT-------KLPKGK   942

Qy   467 DSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLR   526
          || : | | || | | | |: || | :: || ||:: | | ||||||
Db   943 HSVKGLGKTAPDPT--ATTTLDG--VEVPLGNGIS-TGINDTCLLYNEYIVYDVAQVNLK   997

Qy   527 YLLEI   531
          ||||:
Db   998 YLLKL   1002

RESULT 7
T01311
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C;Accession: T01311; S65562
R;Kalicki, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of A. thaliana T14P8.
A;Reference number: Z14290
A;Accession: T01311
A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 1-635 <KAL>
A;Cross-references: EMBL:AF069298; NID:g3193282; PID:g3193299
A;Experimental source: cultivar Columbia
R;Lepiniec, L.; Babiychuk, E.; Kushnir, S.; Van Montagu, M.; Inze, D.
FEBS Lett. 364, 103-108, 1995
A;Title: Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(ADP-r
A;Reference number: S65662; MUID:95269779
A;Accession: S65662
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Cross-references: EMBL:Z48243; NID:g853721; PIDN:CAA88288.1; PID:g853722
C;Genetics:
A;Gene: PARP
A;Map Position: 4
A;Introns: 28/3; 116/1; 162/3; 190/1; 268/3; 312/1; 330/3; 355/2; 371/1; 387/3; 408/3; 4
A;Note: T14P8.19
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger

```
Query Match        24.3%;    Score 683.5;   DB 2;    Length 635;
Best Local Similarity   34.5%;   Pred. No. 1.5e-41;
Matches  190;  Conservative  90;  Mismatches  220;  Indels  51;  Gaps  19;

Qy    2  APKRKASVQTEGSKKQRQTEEEDSFRS-TAEALRAAPADNRVIRVDPSCPFSRNPGIQ 59
Db  109  APVKSSNDEAE--DDNNGEEEKKEEKIVTATKKGAAVLDQWI----PDEIKSQYHVLQ 161

Qy   60  VHED-YDCTLNQTNIGNNNNKFYIIQLLEEGSR-FFCWNRWGRVGEVGQSKMN-HFTCL 115
Db  162  RGDDVYDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVTRWGRVGKGQSKLDGPYDSW 221

Qy  116  EDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPV 175
Db  222  DRAIEIFTNKFNDKTKNYWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNNDIPSSSS-- 278

Qy  176  RTVVKP--CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALE 233
Db  279  --EVPEQSKLDTRVAKFISLICNVSMMAQHMEIGYNANKLPLGKISKSTISKGYEVLK 336

Qy  234  ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPP--INSPDVLQAKKDMLLVLADI 291
Db  337  RISEVIDR--YDRTRLEELSGEFYTVIPHDFGFGFKKMSQFVIDTPQKLKQKIEMVEALGEI 394

Qy  292  ELA-QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---G 347
Db  395  ELATKLLSVDPGLQDD------PLYYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKT 447

Qy  348  NSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH--- 404
Db  448  HSGYTVEIAQLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAP 507

Qy  405  -SGGRVGKGIYFASENSKSAGYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLK 460
Db  508  VTGYMFGKGVYFADMVSKSANY----CYANTGANDGVLLLCEVALGDMNELLYSDYNAD 562

Qy  461  SPPPGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSSFSQSEYLIYKE 520
Db  563  NLPPGKLSTKGVGKTAPNPSEAQTLE-DG--VVPLGKPVE-RSCSKGMLLYNEYIVYNV 618

Qy  521  SQCRLRYLLEI 531
Db  619  EQIKMRYVIQV 629
```

RESULT 8
S04200
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
A;Accession: S04200
R;Huppi, K.; Bhatia, K.; Siwarski, D.; Klinman, D.; Cherney, B.; Smulson, M.

Nucleic Acids Res. 17, 3387-3401, 1989
A;Title: Sequence and organization of the mouse poly (ADP-ribose) polymerase gene.
A;Reference number: S04200; MUID:89263780
A;Accession: S04200
A;Molecule type: DNA
A;Residues: 1-1013 <HUP>
A;Cross-references: EMBL:X14206; NID:g49893; PIDN:CAA32421.1; PID:g49894
C;Genetics:
A;Map position: 1
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc

```
Query Match        24.3%;    Score 683.5;   DB 1;    Length 1013;
Best Local Similarity   32.4%;   Pred. No. 2.9e-41;
Matches  177;  Conservative  105;  Mismatches  215;  Indels  49;  Gaps  17;

Qy    1  MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
Db  493  VAPRGKSAAPSKKSK---GCFKEEGVNKSEKRMKLTLKGAAVDPDSGLEHSAHVLEKG 548

Qy   61  HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
Db  549  GKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRLGTVIGSNKLEQMPSKEE 608

Qy  118  AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
Db  609  AVEQFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKK---------L 654

Qy  178  VVKP---CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA 234
Db  655  TVKPGTKSKLPKPVQELVGMIFDVDSMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSE 714

Qy  235  LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELA 294
Db  715  VQQPVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVA 773

Qy  295  QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSY 350
Db  774  YSLLRGGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEVIRKVKNTHATTHNAY 828

Qy  351  RCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SG 406
Db  829  DL-EVIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTG 887

Qy  407  GRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGF 466
Db  888  YMFGKGIYFADMVSKSANYCHTSQ--GDPIGLIMLGEVALGNMYELK-HASHISKLPKGK 944

Qy  467  DSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPS-FKSSSFSQSEYLIYKESQCRL 525
Db  945  HSVKGLGKTTPDPSASITLE---GVEVPLGTGI-PSGVNDFALLYNEYIVDIAQVNL 998

Qy  526  RYLLEI 531
Db  999  KYLLKL 1004
```

RESULT 9
S26057
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 18-Jun-1999
C;Accession: S26057; S78453; I52331
R;Thibodeau, J.; Gradwohl, G.; Dumas, C.; Clairoux-Moreau, S.; Brunet, G.; Penning, C
Biochem. Cell Biol. 67, 653-660, 1989
A;Title: Cloning of rodent cDNA encoding the poly(ADP-ribose) polymerase catalytic do
A;Reference number: I52331; MUID:90027702
A;Accession: S26057
A;Molecule type: mRNA
A;Residues: 1-500 <THI>
A;Cross-references: EMBL:X65497; NID:g56849; PIDN:CAA46478.1; PID:g56850
R;Potvin, F.

submitted to the EMBL Data Library, March 1992
A;Reference number: S78453
A;Accession: S78453
A;Molecule type: mRNA
A;Residues: 1-124,'H',126-127,'A',129-238,'D',240-500 <POT>
A;Cross-references: EMBL:X65497
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransfer

```
Query Match             24.3%;  Score 682.5; DB 2;  Length 500;
Best Local Similarity   36.0%;  Pred. No. 1.3e-41;
Matches 173; Conservative  86; Mismatches 181; Indels  41;  Gaps  17;

Qy     64 YDCTLNQTNIGNNNKFYIIQLLE--EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKK 120
             :     |:   |||| | :|||||  ::|  |||||||| :|:   ||||
Db     39 FSATLGLVDIVKGTNSYYKLQLLESDKESRYWIFRSWGRVGTVIGSNKLEQMPSKEDAVE  98

Qy    121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK-ALSPQVDSGPVRTVV 179
            ||:| |||  :    : ||: :  ::  : ||| ||  |:||
Db     99 HFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKLAVKPGTKS------- 149

Qy    180 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 239
             |      :   |::||::|   | ||||||  :::|||||::||
Db    150 ---KLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQAV 206

Qy    240 KNPTGDGQSLEELSSCFYTVIPINHFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQA 299
             ::: : ||||| : |:|| |||||:||:||||||:|||:|| ::|| | ||| |:::|
Db    207 SQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNTDSVQAKVEMLDNLLDIEVAYSLLR 265

Qy    300 APGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNL 355
            | ||:   : |   :   ||:    || |::|||||:|||::|       |||| |||
Db    266 GGSDDSSK-----DPIDVNYEKLKTDIKVDRDSEEAEVIRKYVKNTHATTHNAYDL-EV 319

Qy    356 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK 411
             |  ||:    | :|| |||||||||:||||||||||:|:::|||||:      ||||||
Db    320 IDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGK 379

Qy    412 GIYFASENSKSAGYVTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIA 471
            ||||| |:||    || | |||||||:|:| ||||||||||:||||||||||||||:|
Db    380 GIYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSVKG 436

Qy    472 RQGTEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRYLLE 530
             |:: |:|:| :   | ||||| ||| ||||| ||||||||||||||||||||||:::
Db    437 LGKTAPDPSASI--TLDG--VEVPLGTGI--PSGVNDTCLLYNEYIVYDIAQVNLKYLLK 490

Qy    531 I 531
            |
Db    491 L 491
```

RESULT 10
S42208
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - flesh fly (Sarcophaga peregrina)
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Sarcophaga peregrina
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S42208; S71496
R;Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurat
Eur. J. Biochem. 220, 607-614, 1994
A;Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from Sarc
A;Reference number: S42208; MUID:94170813
A;Accession: S42208
A;Molecule type: mRNA
A;Residues: 1-996 <MAS>
A;Cross-references: EMBL:D16482; NID:g473742; PIDN:BAA03943.1; PID:g538248
A;Accession: S71496
A;Molecule type: protein
A;Residues: 170-188;721-736;813-819;879-885 <MAX>
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger
F;1-369/Domain: DNA binding #status predicted <DNA>
F;370-507/Domain: auto-modification #status predicted <AMO>

F;508-996/Domain: NAD binding #status predicted <NAD>

```
Query Match             22.9%;  Score 645.5; DB 1;  Length 996;
Best Local Similarity   32.8%;  Pred. No. 1.6e-38;
Matches 175; Conservative 101; Mismatches 205; Indels  53;  Gaps  18;

Qy     22 EEEDSFRSTAEALRAAPADNRV-----IRVDPSCPF------SRNPGIQVHEDYDCTL  68
           ||| |||||||| |  :| :||     |||| | |       |:|  |||::|: ||
Db    489 EESKSSKSKSIYTKSVPKSMTLKIKDGLAVDPDSGLEDVAHVYVSRN----KEKYNVVL 543

Qy     69 NQTNIGNNNKFYIIQLLEEG--SRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKKK 125
           | ||||||||||||||| |   ||||||||:| :| :| | || ::|  |||||::::
Db    544 GITDIQKNKNSFYKLQLLESDMKNRFWFVFRSWGRIGTTIGGNKLDNFSNLVDAIVQFKEL 603

Qy    126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLD 185
            | | |:|  |::  ::      ::   : |                :|
Db    604 YLEKSGNHFENRENFVKVAGRMYPIDIDYAEDS--------------KIDLSAEHDIKSKLPL- 652

Qy    186 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 245
           |:|  ||:|: ||::|    |::|:||  |||||| ||||||| ||:||:|:|||  |
Db    653 -SVQDIIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLTEIYELIQG-GGT 710

Qy    246 GQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE 305
           ||::|| | |||||||:|||||||| ||||:||||:||:|||||:||| :|:  :
Db    711 NAKFIDATNRFYTLIPHNFGTQSPPLLDTEQVEQLRQMLDSLIEIECAYSLLQT---ED 767

Qy    306 EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKVN 362
           : :  :|  :|   ||: |:|| |:||||||   : ||| |  |:|||| |   | || |
Db    768 SKAD--INPIDKHYEQLKTKLEPLDKNSEEYILLQKYVKNTHAETHKLYDLEVVDIFKVA 825

Qy    363 REGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASE 418
           ||| :|||: ||||:|||||||:  :  |||:|:|||||||      ||| ||||||:|
Db    826 RQGEARRYKPFKKLHNRRLLWHGSRLTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADM 885

Qy    419 NSKSAGYVTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPD 478
           :||| |:| | |||||||:|:|||||||||||||||||||:|||||||||| |::||:|
Db    886 VSKSANYCCTSH--HNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFGRGRTMPN 942

Qy    479 PAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIH 532
           |::  :   | || :    |:|||| |||||||||||||||||||||||:::|
Db    943 PSESIIRE-DG--VEIPLGKPITNDSLK-SSLLYNEFIIYDIAQVNIQYMLRMN 992
```

RESULT 11
T51353
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) [imported] - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C;Accession: T51353
R;Doucet-chabeaud, G.; Kazmaier, M.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z25379
A;Accession: T51353
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-983 <DOU>
A;Cross-references: EMBL:AJ131705; PIDN:CAA10482.1
A;Experimental source: cultivar landsberg erecta
C;Genetics:
C;Gene: parp-1
C;Function:
A;Description: ADP-ribose polymer synthesis
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

```
Query Match             22.1%;  Score 621.5; DB 2;  Length 983;
Best Local Similarity   32.7%;  Pred. No. 8.8e-37;
Matches 179; Conservative  87; Mismatches 209; Indels  73;  Gaps  21;

Qy     15 KKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH----ED----YDC  66
```

```
Db      472 KKQRKLPFDKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT 524
Db  551 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGGNKVEEMS-KSDAVHEF 609
Qy  123 KKKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVV 179
Db  610 KRLFLEKTGNTWESWEQKTNFQKOPGKFLPLDIDYGVNKQVA---------KKEPFQT-- 658
Qy  180 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL----EAL 235
Db  659 -SSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQRLL 717
Qy  236 EEAMKNPTGDGQSLEELSSCFYTVIPIHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQ 295
Db  718 TESDPQPTMKESLLVDASNRFFTMIP----SIHPHIIRDEDDFKSKVKMLEALQDIEIAS 773
Qy  296 TLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP-- 353
Db  774 RI---VGFDVDSTES---LDDKYKKLHCDISPLPHDSEDYRLIEKYL----NTTHAPTH 822
Qy  354 -----NLRHVWKVNREGEGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRI----MP 403
Db  823 TEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAP 882
Qy  404 HSGGRVGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPP 463
Db  883 ATGYMFGKGIYFADLVSKSAQYCYT--CKKNPVGLMLLSEVALGEIHELT-KAKYMDKPP 939
Qy  464 PGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPVVQCPSFKSSSFSQSEYLIYKESQC 523
Db  940 RGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPVS-SKVKASELMYNEYIVYDTAQV 995

Qy  524 RLRYLLEI 531
Db  996 KLQFLLKV 1003
```

RESULT 13
A47474
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: A47474
R;Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.
Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993
A;Title: Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase: leucine zip
A;Reference number: A47474; MUID:93234521
A;Accession: A47474
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-994 <UCH>
A;Cross-references: GB:D13806; GB:D13807; GB:D13808; NID:g303545; PIDN:BAA02964.1; PI
A;Note: sequence extracted from NCBI backbone (NCBIN:129703, NCBIP:129704)
C;Genetics:
A;Gene: FlyBase:Parp
A;Cross-references: FlyBase:FBgn0010247
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match   22.0%;   Score 619; DB 1; Length 994;
Best Local Similarity   31.5%;   Pred. No. 1.4e-36;
Matches 168;   Conservative 100; Mismatches 222;   Indels 44; Gaps 17;

```
Qy   16 KQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED----YDCT 67
Db  481 KSRIPKETTKSLNSNSIYTKSMPV-SRTFKVKDGLAVDPDSGLEDIAHVYVDSNNKIYSVV 539
Qy   68 LNQTNIGNNNKFYIIQLL--EEGSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKK 124
Db  540 LGLTDIQRNKNSYYKVQLLKADKKEKYWIFRSWGRIGNSKLEEFDTSESAKRNFKE 599
Qy  125 KFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVKPCSL 184
Db  600 IYADKTGNEYEQRDNFVKRTGRMYPIEIQ--YDDDQKLVKHESHFFTS--------KL 647
```

```
Db      472 KKQRKLPFDKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT 524
Qy   67 TLNQTNIGNNNKFYIIQLLEE--GSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDF 122
Db  525 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGGNKVEEMS-KSDAVHEF 583
Qy  123 KKKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVV 179
Db  584 KRLFLEKTGNTWESWEQKTNFQKOPGKFLPLDIDYGVNKQVA---------KKEPFQT-- 632
Qy  180 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL----EAL 235
Db  633 -SSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQRLL 691
Qy  236 EEAMKNPTGDGQSLEELSSCFYTVIPIHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQ 295
Db  692 TESDPQPTMKESLLVDASNRFFTMIP----SIHPHIIRDEDDFKSKVKMLEALQDIEIAS 747
Qy  296 TLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP-- 353
Db  748 RI---VGFDVDSTES---LDDKYKKLHCDISPLPHDSEDYRLIEKYL----NTTHAPTH 796
Qy  354 -----NLRHVWKVNREGEGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRI----MP 403
Db  797 TEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAP 856
Qy  404 HSGGRVGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPP 463
Db  857 ATGYMFGKGIYFADLVSKSAQYCYT--CKKNPVGLMLLSEVALGEIHELT-KAKYMDKPP 913
Qy  464 PGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPVVQCPSFKSSSFSQSEYLIYKESQC 523
Db  914 RGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPVS-SKVKASELMYNEYIVYDTAQV 969

Qy  524 RLRYLLEI 531
Db  970 KLQFLLKV 977
```

RESULT 12
C84719
probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: C84719
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: C84719
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1009 <STO>
A;Cross-references: GB:AE002093; NID:g4432827; PIDN:AAD20677.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g31320
A;Map position: 2
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match   22.1%;   Score 621.5; DB 2; Length 1009;
Best Local Similarity   32.7%;   Pred. No. 9.1e-37;
Matches 179;   Conservative 87; Mismatches 209; Indels 73; Gaps 21;

```
Qy   15 KKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH----ED----YDC 66
Db  498 KKQRKLPFDKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT 550
Qy   67 TLNQTNIGNNNKFYIIQLLEE--GSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDF 122
```

```
Qy  185  DPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTG  244
Db  648  EISVQNLIKLIFIDSMNKTLMEFHIDMDKMPLGKLSAHQIQSAYRVVKEIYNVLECGSN   707

Qy  245  DGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEE   304
Db  708  TAK-LIDATNRFYTLIPHNFGVQLPTLIETHQQIEDLRQMLDSLAEIEVAYSI-----IK  761

Qy  305  EEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKV  361
Db  762  SEDVSDACNPLDNHYAQIKTQLVALDKNSEEFSILSQIVKNTHASTHKSYDLKIVDVFKV  821

Qy  362  NREGEGDRFQAHSKLGNRRLLWHGTNVAVVAALTSGLRI---MPHSGGRVGKGIYFAS    417
Db  822  SRQGEARRFKPFKKLHNRKLLWHGSRLTNFVGILSHGLRIAPPEAPTGYMFGKGIYFAD   881

Qy  418  ENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEP  477
Db  882  MVSKSSANYCCTSQ-QNSTGLMLLSEVALGDMMECT-SAKYINKLSNNKHSCFGRGRTMP  938

Qy  478  DPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI        531
Db  939  DPTKSY-IRSDG--VEIPYGETITDEHLK-SSLLYNEYIVYDVAQVNIQYLFRM        988
```

RESULT 14
T03657
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) 2 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
C;Accession: T03657
R;Babiychuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Monta
submitted to the EMBL Data Library, November 1997
A;Reference number: Z14992
A;Accession: T03657
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-969 <BAB>
A;Cross-references: EMBL:AJ222589; PIDN:CAA10889.1
C;Genetics:
A;Gene: PARP2
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

```
Query Match             20.8%;   Score 584.5;  DB 2;   Length 969;
Best Local Similarity   30.7%;   Pred. No. 4.1e-34;
Matches   164;  Conservative   98;  Mismatches   199;  Indels   73;  Gaps   20;

Qy   33  ALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED----YDCTLNQTNIGNNNNKFYIIQ   84
Db  468  ALESSKGSTVTVKVKGRSAVHESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQ  527

Qy   85  LLE-EGSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDFKKKFWEKTKNKWEE--R   137
Db  528  IIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEKTGNSWEAWECK  586

Qy  138  DRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLDPATQNLITNIFS  197
Db  587  TNFRKQPGRFYIPLDV------DYGVKKAPKRKDISEMKS-----SLAPQLLELMKMLFN  634

Qy  198  KEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSLEE-----L  252
Db  635  VETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAA  694

Qy  253  SSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEEKVEEVP   312
Db  695  SNRFFTLIP----SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDES-----  745

Qy  313  HPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------NLRHVWKVNREG   365
Db  746  --LDDKYMKLHCDITPLAHDSEDYKLIEQYLLNT----HAPTHKDWSLELEEVFSLDRDG  799

Qy  366  EGDRFQAH-SKLGNRRLLWHGTNVAVVAALTSGLRIMPH----SGGRVGKGIYFASENS  420
Db  800  ELNKYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVS  859

Qy  421  KSAGYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEP  477
Db  860  KSAQY-----CYVDRNNPVGLMLLSEVALGDMYELK-KATSMDKPPRGKHSTKGLGKTVP  913

Qy  478  DPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI        531
Db  914  LESEFVKWRDD---VVVPCGKPVP-SSIRSSELMYNEYIVNTSQVKMQFLLKV         963
```

RESULT 15
T18600
hypothetical protein AC8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18600
R;McMurray, A.;
submitted to the EMBL Data Library, November 1996
A;Reference number: Z18996
A;Accession: T18600
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-727 <WIL>
A;Cross-references: EMBL:Z83097; PIDN:CAB05448.1; GSPDB:GN00023; CESP:AC8.1
A;Experimental source: clone AC8
C;Genetics:
A;Gene: CESP:AC8.1
A;Map position: 5
A;Introns: 31/3; 271/3; 312/3; 348/1; 466/3; 545/3; 617/3; 702/2

```
Query Match             16.9%;   Score 475.5;  DB 2;   Length 727;
Best Local Similarity   30.2%;   Pred. No. 2.2e-26;
Matches   134;  Conservative   68;  Mismatches   170;  Indels   71;  Gaps   13;

Qy  126  FWEKTKNKWEERDRFVAQPNKYTLIEVQGE--AESQEAVVKALSPQVDSGPVRTVVKPCS  183
Db  315  FHEKTKNDWIYRKHFRKMPGMFSYVETDYSEFVGTNNGHKKKITPGSKITPGSKTLLPKS  374

Qy  184  LDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEEAMKNPT    243
Db  375  V-----KEVVMSIFDVENMKSALKSFEIDVNKMPLGRLSHNQINLAFEVLNDISDLLVKLP 430

Qy  244  GDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGE  303
Db  431  IDASKILDFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNMLNALLDIKFAYD-QISGGD  489

Qy  304  EEEKVEEVPHPLDRDYQLLRCQLQLLDSGSEYKAIQTYLKQT-GNSYRCP-NLRHVWKV  361
Db  490  VPASTSLSIDPVDINYRKLKCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVELIDILKV  549

Qy  362  NREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFAS  417
Db  550  NRDNESSSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPPEAPVSGYMFGKGVYFAD 607

Qy  418  ENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFDSVIARGQTE  476
Db  608  MFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSV-------  657

Qy  477  PDPAQDIELELDGQPVVVPQGPPVQCP----SFKSS--------SFSQ-----------  512
Db  658  ------------------QGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQGKQNVDYHL   698

Qy  513  --SEYLIYKESQCRLRYLLEIHL     533
Db  699  LYNEFIVYDVDQIQLKYLVRVKM     721
```

us-09-701-586b-8.rpr

Search completed: August 29, 2002, 07:59:37
Job time: 284 sec

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 29, 2002, 07:59:37 ; Search time 69.02 Seconds
                                         (without alignments)
                                        735.079 Million cell updates/sec

Title:           US-09-701-586B-10
Perfect score:   2789
Sequence:        1 MAPKRRASVQTEGSKKQRQG..........EXYLIYKESQCRLRYLLEIHL 528

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters:      283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_71:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1987.5	71.3	459	2	T08713	NAD+ ADP-ribosyltr
2	719.5	25.8	653	2	T03656	probable NAD+ ADP-
3	707.5	25.4	1014	1	A29725	NAD+ ADP-ribosyltr
4	704.5	25.4	1016	1	JS0428	NAD+ ADP-ribosyltr
5	691	24.8	635	2	T01311	NAD+ ADP-ribosyltr
6	689	24.7	1011	1	JH0581	NAD+ ADP-ribosyltr
7	688	24.7	1013	1	S04200	NAD+ ADP-ribosyltr
8	686	24.6	500	2	S26057	NAD+ ADP-ribosyltr
9	681	24.4	998	2	S31735	NAD+ ADP-ribosyltr
10	652	23.4	996	2	S42208	NAD+ ADP-ribosyltr
11	624	22.4	983	2	T51353	NAD+ ADP-ribosyltr
12	624	22.4	1009	2	C84719	probable poly (ADP
13	621.5	22.3	994	1	A47474	NAD+ ADP-ribosyltr
14	584	20.9	969	2	T03657	NAD+ ADP-ribosyltr
15	481	17.2	727	2	T18600	hypothetical prote
16	389	13.9	538	2	T20414	hypothetical prote
17	193	6.9	135	2	PN0494	NAD+ ADP-ribosyltr
18	172.5	6.2	2004	2	D88948	protein ZK1005.1
19	121	4.4	1223	2	T17345	hypothetical prote
20	121.5	4.3	1092	2	T37317	carbamoyl-phosphat
21	120.5	4.3	181	2	T03058	NAD+ ADP-ribosyltr
22	120	4.3	5175	2	T20992	hypothetical prote
23	120	4.3	5198	2	T43290	hemicentin precurs
24	118	4.2	935	2	E96806	hypothetical prote
25	117	4.2	954	2	E86174	protein F19P19.26
26	112	4.0	1203	2	D86625	subunit A of ATP-d
27	111.5	4.0	359	2	T12540	hypothetical prote
28	107.5	3.9	638	2	C83503	exotoxin A precurs
29	106.5	3.8	646	2	T38022	probable GTP-bindi
30	105.5	3.8	3075	2	S14458	laminin alpha-1 ch
31	104	3.7	2285	2	T12796	probable transglyc
32	103.5	3.7	638	2	A30347	exotoxin A precurs
33	103	3.7	357	1	QQMSLL	retrovirus-related
34	103	3.7	1031	1	A38713	kinesin heavy chai
35	103	3.7	1435	2	S69632	regulatory protein
36	102.5	3.7	354	2	D90187	oxidoreductase (im
37	102.5	3.7	612	1	WDBPT5	gene D13 protein -
38	102.5	3.7	2484	2	T26216	hypothetical prote
39	102.5	3.7	2607	2	T26215	hypothetical prote
40	101.5	3.6	295	2	F83412	probable transcrip
41	101.5	3.6	398	2	E71424	hypothetical prote
42	101.5	3.6	424	2	H85175	hypothetical prote
43	101.5	3.6	436	2	C90043	IgG-binding protei
44	101.5	3.6	769	2	E96613	hypothetical prote
45	101.5	3.6	1108	2	E71104	probable cell divi

ALIGNMENTS

```
RESULT   1
T08713
NAD+ ADP-ribosyltransferase homolog DKFZp566G0224.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Nov-2000
C;Accession: T08713
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.;
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16472
A;Accession: T08713
A;Molecule type: mRNA
A;Residues: 1-459 <ANS>
A;Experimental source: fetal kidney; clone DKFZp566G0224
C;Genetics:
A;Note: DKFZp566G0224.1
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match            71.3%;   Score 1987.5;   DB 2;   Length 459;
Best Local Similarity  82.4%;   Pred. No. 9e-136;
Matches  379;  Conservative  32;  Mismatches  44;  Indels  5;  Gaps  3;

Qy    73 IGNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLEDAKKDFKKKFWEKTK 131
            |||||||||:|||||||:|| |||||:||||||||||:||||:||||||||||:|||:|
Db     1 IENNNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREKTK  60

Qy   132 NKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVKVDSGPVRTV---VKPCSLDPATQNLIT 188
            | | ||| |||| ||||||||||||||:||||||:|||| ||||:|  ||||||||||:| |||
Db    61 NNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVKVDRGPVRTVTKRVQPCSLDPATQKLIT 120

Qy   189 NIFSKEMPKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSLEELS 248
            |||||||:|:||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTGDGGQSLEELS 180

Qy   249 SCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVEEVPH 308
            ||| ||||||||:||||||||||||||||||||||||||||:||||||||:||:|:||||||
Db   181 SHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAV-SEQEKTVEEVPH 239

Qy   309 PLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAH 368
            |||||||||||:||||||||||||:||:||||||||:|:||||||:||||:||||||||||||
Db   240 PLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVNQEGEEDRFQAH 299

Qy   369 SKLGNRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGG 428
            |||||:|||||||:||||||||||||||||||||||||||||||||||||||||:||||||
Db   300 SKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGEGIYFASENSKSAGYVIGMKCGA 359

Qy   429 HQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPVV 488
            | |:||||||||| || ||:|||:|||||||||||||||||:||||||||||:|||||||||
Db   360 HHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVV 419

Qy   489 VPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
```

```
              ||||  |||  ||  ||:|||||||||::||||||||||:||
Db   420 VPQGQPVPCPEFSSTFSQSEYLIYQESQCRLRYLLEVHL 459


RESULT    2
T03656
probable NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Oct-1999
C;Accession: T03656
R;Babyichuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Monta
submitted to the EMBL Data Library, November 1997
A;Description: Higher plants possess two poly(ADP-ribose) polymerases.
A;Reference number: Z14991
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-653 <BAB>
A;Cross-references: EMBL:AJ222588; NID:e1264090; PIDN:CAA10888.1; PID:e1264091
C;Genetics:
A;Gene: PARP
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match             25.8%;  Score 719.5;  DB 2;  Length 653;
Best Local Similarity   35.3%;  Pred. No. 5.5e-44;
Matches 190;  Conservative 91;  Mismatches 205;  Indels 53;  Gaps 17;

Qy    13 GSKKQRQGTEEEDSFRS----TAEALRAAPADNRV---IRVDPSCPFSRNPGIQVHED-Y  64
            :|: :: :|           |||:  |     ||| :::  |
Db   138 GAKEVIKGGDEEVEVKKEKMVTATKKGAAVLDQHIPDHIKVNYHV------LQVGDEIY  190

Qy    65 DCTLNQTNIGNNNNKFYIIQLLEE--GSRFFCWNRWGRVGEVGQSKMNHFTCLED-AKKD 121
            ||||||||||||||||:|:||      |||||| | ||||| |
Db   191 DATLNQTNVGDNNNKFYIIQVLESDAGGSFMVYNRWGRVGRGDKLHGPSPTRDQAIYE   250

Qy   122 FKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQ-GEAEESQEAVVKVDSGPVRTVVKPCSLD 180
           |: |   |||   : ||| | :: |||
Db   251 FEGKFPHNKTNNHWSDRKNFKCYAKKYTWLEMDYGETEKE----IEKGSITDQIKETKLE  305

Qy   181 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAEMKNPTGD    240
           | |:::||:|:|  ||:|  | :|| || |  :|   |:   |:  |
Db   306 TRIAQFISLICNISMMKQRMVEIGYNAEKLPLGKLRKATILKGYVHVLKRISDVISK--AD 363

Qy   241 GQSLEELSSCFYTVIPHFG--RSRPPPINSPDVLQAKKDMLLVLADIELAQTIQAAPGE   298
           |:|:||||:|||  |||| ||| |:| |:   |||| | |  :
Db   364 RRHLEQLTGEFYTVIPHDFGFRKMREFIIDTPQKIKAKLEMVEALGEIFIATKLL-----  418

Qy   299 EEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRCPNLRHVWK  355
           :: |  |  :|  ::|| |  || |   : ||:|  :|   |:   :|:
Db   419 -EDDSSDQDDPLYARYKQLHCDFTPLEADSDEYSMIKSYLRNTHGKTHSSYTVDIVQIFK  477

Qy   356 VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH--SGGRVGKGIYFA    411
           |:     | : :  ||| |||  ||      |  |:  || |      |||| |:|:
Db   478 VSRHGETERFQKFASTRNRMLLWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGKGVIFA  537

Qy   412 SENKSAGYVTTMHCGGHQV--GYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG   468
            ||||  ||||||:|| |  |:||| |||| |  :|  || :|
Db   538 DMFSKSANY----CYASEACRSGVLLCEVALGDMNELLNADYDANNLPKGKLRSKGVG   592

Qy   469 QTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIH   527
            ||:  |   ::|::  ||| :|     | || |  |::    |:::|||||||
Db   593 QTAPNM---VESKVADDGVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMRVVLHVN   647


RESULT    3
A29725
NAD+ ADP-ribosyltransferase (EC 2.4.2.30), nuclear - human
N;Alternate names: poly (ADP-ribose) polymerase; poly (ADP-ribose) synthetase; poly(ADP)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29725; A28498; A39976; A26901; I38096; B33321; A33321; A35635; A61559; S14
R;Uchida, K.; Morita, T.; Sato, T.; Ogura, T.; Yamashita, R.; Noguchi, S.; Suzuki, H.; N
```

```
Biochem. Biophys. Res. Commun. 148, 617-622, 1987
A;Title: Nucleotide sequence of a full-length cDNA for human fibroblast poly(ADP-ribo
A;Reference number: A29725; MUID:88076933
A;Accession: A29725
A;Molecule type: mRNA
A;Residues: 1-69,'Q',71-1014 <UCH>
A;Cross-references: GB:M18112; NID:g190166; PIDN:AAA60137.1; PID:g190167
R;Kurosaki, T.; Ushiro, H.; Mitsuuchi, Y.; Suzuki, S.; Matsuda, M.; Matsuda, Y.; Katu
J. Biol. Chem. 262, 15990-15997, 1987
A;Title: Primary structure of human poly (ADP-ribose) synthetase as deduced from cDNA
A;Reference number: A28498; MUID:88058958
A;Accession: A28498
A;Molecule type: mRNA
A;Residues: 1-16,'E',18-211,'K',213-236,'R',238-366,'H',369-1014 <KUR>
A;Cross-references: GB:J03473
R;Cherney, B.W.; McBride, O.W.; Chen, D.; Alkhatib, H.; Bhatia, K.; Hensley, P.; Smul
Proc. Natl. Acad. Sci. U.S.A. 84, 8370-8374, 1987
A;Title: cDNA sequence, protein structure, and chromosomal location of the human gene
A;Reference number: A39976; MUID:88068596
A;Accession: A39976
A;Molecule type: mRNA
A;Residues: 1-49,'D',51-612,'Q',614-907,'Y',909-939,'R',941-979,'I',981-1014 <CHE>
A;Cross-references: GB:J03030
A;Note: the authors translated the codon ATA for residue 980 as Asn
R;Suzuki, H.; Uchida, K.; Shima, H.; Sato, T.; Okamoto, T.; Kimura, T.; Miwa, M.
Biochem. Biophys. Res. Commun. 146, 403-409, 1987
A;Title: Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and expressi
A;Reference number: A26901; MUID:87298455
A;Accession: A26901
A;Molecule type: mRNA
A;Residues: 441-610,'N',612-880;921-1014 <SUZ>
A;Note: the sequence figure has an omission of forty residues
R;Ogura, T.; Nyunoya, H.; Takahashi-Masutani, M.; Miwa, M.; Sugimura, T.; Esumi, H.
Biochem. Biophys. Res. Commun. 167, 701-710, 1990
A;Title: Characterization of a putative promoter region of the human poly(ADP-ribose)
A;Reference number: I38096; MUID:90211250
A;Accession: I38096
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-40 <RES>
A;Cross-references: EMBL:X16674; NID:g510112; PIDN:CAA34663.1; PID:g1017423
R;Auer, B.; Nagl, U.; Herzog, H.; Schneider, R.; Schweiger, M.
DNA 8, 575-580, 1989
A;Title: Human nuclear NAD(+) ADP-ribosyltransferase(polymerizing): organization of t
A;Reference number: A33321; MUID:90091744
A;Accession: A33321
A;Molecule type: DNA
A;Residues: 38-43;93-98;132-137;204-209;237-242;276-281;335-340;384-389;431-436;512-5
A;Cross-references: GB:M29544; GB:M22953
A;Note: the authors translated the codon GTG for residue 54 as Glu
A;Note: these fragments represent intron-exon boundaries
A;Accession: A33321
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 16-66;96;121-159,'D',161-167 <AU2>
A;Note: these fragments represent a zinc finger-containing DNA-binding region
R;Gradwohl, G.; Menissier de Murcia, J.; Molinete, M.; Simonin, F.; Koken, M.; Hoeijm
Proc. Natl. Acad. Sci. U.S.A. 87, 2990-2994, 1990
A;Title: The second zinc-finger domain of poly(ADP-ribose) polymerase determines spec
A;Reference number: A35635; MUID:90222155
A;Accession: A35635
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 12-26,'T',28-66;116-166 <GRA>
R;Schneider, R.; Auer, B.; Kuehne, C.; Herzog, H.; Klocker, H.; Burtscher, H.J.; Hirs
Eur. J. Cell Biol. 44, 302-307, 1987
A;Title: Isolation of a cDNA clone for human NAD (+): protein ADP-ribosyltransferase.
A;Reference number: A61559; MUID:88082900
A;Accession: A61559
A;Molecule type: mRNA
A;Residues: 381-420;682-710 <SCH>
R;Yokoyama, Y.; Kawamoto, T.; Mitsuuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Tera
Eur. J. Biochem. 194, 521-526, 1990
```

A;Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.
A;Reference number: S14010; MUID:91099327
A;Accession: S14010
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-95 <YOK>
A;Cross-references: EMBL:X56140; NID:g35286; PIDN:CAA39606.1; PID:g825702
C;Comment: This protein can ADP-ribosylate itself as well as other proteins.
C;Genetics:
A;Gene: GDB:ADPRT; PPOL
A;Cross-references: GDB:119508; OMIM:173870
A;Map position: 1q41-1q42
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransfer

```
Query Match              25.4%;   Score 707.5;  DB 1;   Length 1014;
Best Local Similarity    34.0%;   Pred. No. 7.6e-43;
Matches  183;  Conservative  101;  Mismatches  216;  Indels  39;  Gaps  17;

Qy    1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV   60
          :: : || |  ::   : :   ::        :  |     |||| | |: :  :
Db  493 VAPRGKSGA---ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDDPDSGLEHSAHVLEKG 549

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED  117
          ||  | |:|||||||||  :| ||||||    ||: |||||| |: |: |    ::
Db  550 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKED 609

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVK  175
          | : : | : ||| | | || :   | |:| |:| |::: :||||| |     | |
Db  610 AIEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKLTVNPGTKSKLPK 667

Qy  176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK  235
          |      |:| : ::|  |: | |  ||  :| ||||||| | || :: : | :|
Db  668 P------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQAVS 721

Qy  236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTIQAA   295
          : : || || :|| |||:||| ||  :|| :| | ||||::|| |:||||: |:
Db  722 QGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLRG 780

Qy  296 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLR   351
          |: ::: ||  ||:::| :  | : :| ::| :|| : : ::||     : :|  |::
Db  781 GSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATTHNAYDL-EVI 834

Qy  352 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKG  407
          :: |::|||| :|: : :|::|||||||: : ::| |||:|||| ||     ||  :||
Db  835 DIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKG 894

Qy  408 IYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIAR  467
          |||| : ||:| : :: |   ||| | :|||||| ::   : |:  |: : ||  |: :
Db  895 IYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSVKGL 951

Qy  468 GQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI  526
          |:| ||| :: | |   | ||:| :       |:  :|| ||  :| ||||  || |||:
Db  952 GKTTPDPSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEYIVYDIAQVNLKYLLKL 1005
```

RESULT 4
JS0428
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine
N;Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate rib
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: JS0428; S00328; A30458
R;Saito, I.
submitted to JIPID, February 1990
A;Reference number: JS0428
A;Accession: JS0428
A;Molecule type: mRNA
A;Residues: 1-1016 <SAI>
A;Experimental source: thymus
R;Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; Taka
Eur. J. Biochem. 171, 571-575, 1988

A;Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int
A;Reference number: S00328; MUID:88151954
A;Accession: S00328
A;Molecule type: mRNA
A;Residues: 648-714;838-904 <TAN>
A;Cross-references: EMBL:X06986
A;Accession: A30458
A;Molecule type: protein
A;Residues: 658-685;689-696;893-901 <TA2>
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-loo
F;21-51/Region: zinc finger
F;128-165/Region: zinc finger
F;200-220/Region: nuclear location signal
F;224-231/Region: helix-turn-helix motif
F;250-270/Region: helix-turn-helix motif
F;494-501/Region: nucleotide-binding motif A (P-loop)
F;890-903/Region: nucleotide binding #status predicted

```
Query Match              25.3%;   Score 704.5;  DB 1;   Length 1016;
Best Local Similarity    34.4%;   Pred. No. 1.3e-42;
Matches  187;  Conservative  95;  Mismatches  213;  Indels  213;  Gaps  20;

Qy    1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAE--ALRAAPADNRVIRVDPSCPFSRNPG   57
          : | | ||   || :  : | ::  : :       ||:|  |||| |     :  |
Db  495 VGPKGKSGAAP--SKKSKGPVKEEGTNKSEKRMKLTLKGGAA-----VDPDSGLEHNAH 546

Qy   58 I--QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHF  112
            :|  ||  | |:|||||||||  :| ||||||    ||: |||||| |: |: |
Db  547 VLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVIGSNKLEQM 606

Qy  113 TCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK-VDSGPV  170
          : | :| : : | : ||| | | || :   | |:| |:| |::: :||||| |
Db  607 PSKEDAIEHFMKLYEEKTGNAWHSKN-FTKHPKKFYPLEID-YGQDEEAVKKLTVNPGTK 664

Qy  171 RTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAL  230
             |   |    |:| : ::|  |: | |  ||  :| ||||||| | || :: : | :|
Db  665 SKLPKP-----VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEV 718

Qy  231 EEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQ   290
          :| : : : || || :|| |||:||| ||  :|| :| | ||||::|| |:||||:
Db  719 QQALSQGSGSDSHIL-DLSNRFYTLIPHDFGMKKPPLLNNANSVQAKVEMLDNLLDIEVAY 777

Qy  291 TLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYR   346
           :  ||: ::: ||  ||:::| :  | : :| ::| :|| : : ::||     : :|
Db  778 SLLRGGSDDSSK----DPIDVNYEKLKTDIKVVDKDSEEAEIIRKYVKNTHATTHNAYD 832

Qy  347 CPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGG  402
            :: |::|||| :|: : :|::|||||||: : ::| |||:||||  ||     ||
Db  833 L-EVVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTNFAGILSQGLRIAPPEAPVTGY 891

Qy  403 RVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFD  462
           :||||||| : ||:| : :: |   ||| | :||| ||| ::   : |:  |: : ||
Db  892 MFGKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEAALGNMYELK-HARHISKLPKGKH 948

Qy  463 SVIARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRY  522
          ||  |||: ||| :: | |   | ||:| :       |:  :|| ||  :| ||||   |
Db  949 SVKGLGKTTPDPSASI--TVDG--VEVPLGTGIS-SGVNDTCLLYNEYIVYDIAQVHLKY 1003

Qy  523 LEI  526
          |
Db 1004 LLKL 1007
```

RESULT 5
T01311
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C;Accession: T01311; S65662
R;Kalicki, J.; Elliott, G.; Cloud, J.

```
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of A. thaliana T14P8.
A;Reference number: Z14290
A;Accession: T01311
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-635 <KAL>
A;Cross-references: EMBL:AF069298; NID:g3193282; PID:g3193299
A;Experimental source: cultivar Columbia
R;Lepiniec, L.; Babiychuk, E.; Kushnir, S.; Van Montagu, M.; Inze, D.
FEBS Lett. 364, 103-108, 1995
A;Title: Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(ADP-r
A;Reference number: S65662; MUID:95269779
A;Accession: S65662
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-115,'GT',116-635 <LEP>
A;Cross-references: EMBL:Z48243; NID:g853721; PIDN:CAA88288.1; PID:g853722
C;Genetics:
A;Gene: PARP
A;Map position: 4
A;Introns: 28/3; 116/1; 162/3; 190/1; 268/3; 312/1; 330/3; 355/2; 371/1; 387/3; 408/3; 4
A;Note: T14P8.19
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger


Query Match           24.8%;  Score 691;  DB 2;  Length 635;
Best Local Similarity 34.6%;  Pred. No. 6.1e-42;
Matches  189;  Conservative  91;  Mismatches  220;  Indels  46;  Gaps  18;

Qy    2 APKRKASVQTEGSKKQRQGTEEEDSFRS--TAEALRAAPADNRVIRVDPSCPFSRNPGIQ 59
Db  109 APVKSSNDEAE---DDNNGFEEEKKEEKIVTATKKGAAVLDQWI----PDEIKSQYHVLQ 161

Qy   60 VHED-YDCTLNQTNIGNNNNKFYIIQLLEEGSR--FFCWNRWGREVGQEVGQSKMN-HFTCL 115
Db  162 RGDDVYDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVTRWGRVGKGQSKLDGPYDSW 221

Qy  116 EDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVK 175
Db  222 DRAIEIFTNKFNDKTKNYWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNNDIFSSSSEVK 281

Qy  176 P--CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEA 233
Db  282 PEQSKLDTRVAKFISLICNVSMMAQHMEIGYNANKLPLGKISKSTISKGYEVLKRISEV 341

Qy  234 MKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPP--INSPDVLQAKKDMLLVLADIELA-Q 290
Db  342 IDR--YDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATK 399

Qy  291 TLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRC 347
Db  400 LLSVDPGLQDD------PLYYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYT 452

Qy  348 PNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGSRTNWAGILSQGLRIMPH----SGGR 403
Db  453 VEIAQLFRASRAVEADRFQOFSSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYM 512

Qy  404 VGKGIYFASENSKSAGYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPG 460
Db  513 FGKGVYFADMFSKSANY------CYANTGANDGVLLCEVALGDMNELLYSDYNADNLPPG 567

Qy  461 FDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRL 520
Db  568 KLSTKGVGKTAPNPSEAQTLE-DG--VVVPLGKPVE-RSCSKGMLLYNEYIVYNVEQIKM 623

Qy  521 RYLLEI 526
Db  624 RYVIQV 629


RESULT 6
JH0581
```

```
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - chicken
N;Alternate names: poly(ADP-ribose) synthase
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JH0581
R;Ittel, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.
Gene 102, 157-164, 1991
A;Title: Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence an
A;Reference number: JH0581; MUID:91340148
A;Accession: JH0581
A;Molecule type: mRNA
A;Residues: 1-1011 <ITT>
A;Cross-references: EMBL:X52690; NID:g63742; PIDN:CAA36917.1; PID:g63743
C;Comment: This protein is a chromatin-bound enzyme.
C;Comment: This enzyme catalyzes DNA-dependent post-translational modifications of va
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc


Query Match           24.7%;  Score 689;  DB 1;  Length 1011;
Best Local Similarity 33.9%;  Pred. No. 1.7e-41;
Matches  184;  Conservative  97;  Mismatches  205;  Indels  56;  Gaps  20;

Qy    5 RKASVQTEGSKKQRQG-TEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED----- 63
Db  497 KPANMKSAGKVKEEQGPSKSEKKMKLTVKGGAAVDPDSGL--EDSAHVFEKGGKI----- 549

Qy   64 YDCTLNQTNIGNNNNKFYIIQLLEEG--SRRFCWNRWGRVGEV-GQSKMNHFTCLEDAKK 120
Db  550 FSATLGLVDIVKGTNSYYKLQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAVE 609

Qy  121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVKPCS 178
Db  610 HFLNLYEEKTGNSWHSKN-FTKYPKKFYPLEID-YGQDEEAVRKLTVSAGTKSKLAKP-- 665

Qy  179 LDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPT 238
Db  666 ----IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILNEVQQAVSDGG 721

Qy  239 GDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGE 298
Db  722 SESQIL-DLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRGGNE 780

Qy  299 EEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLRHVW 354
Db  781 DGDK------DPIDINYEKLRTDIKVVDKDSEEAKIIKQYVKNTHAATHNAYDL-KVVEIF 834

Qy  355 KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYF 410
Db  835 RIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYF 894

Qy  411 ASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEH-----HITIDDPSLKSPPPGFDSV 464
Db  895 RDMVSKSANYCHTSQ--ADPIGLILGEVALGNMYELKNASHIT------KLPKGKHSV 945

Qy  465 IARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLYLL 524
Db  946 KGLGKTAPDPT--ATTTLDG--VEVPLGNGIS-TGINDTCLLYNEYIVYDVAQVNLKYLL 1000

Qy  525 EI 526
Db 1001 KL 1002


RESULT 7
S04200
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - mouse
C;Species: Mus musculus (house mouse)
C;Accession: S04200
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
R;Huppi, K.; Bhatia, K.; Siwarski, D.; Klinman, D.; Cherney, B.; Smulson, M.
Nucleic Acids Res. 17, 3387-3401, 1989
A;Title: Sequence and organization of the mouse poly (ADP-ribose) polymerase gene.
```

A;Reference number: S04200
A;Accession: S04200; MUID:89263780
A;Molecule type: DNA
A;Residues: 1-1013 <HUP>
A;Cross-references: EMBL:X14206; NID:g49893; PIDN:CAA32421.1; PID:g49894
C;Genetics:
A;Map position: 1
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc fi

Query Match   24.7%;   Score 688; DB 1; Length 1013;
Best Local Similarity 32.7%; Pred. No. 2e-41;
Matches 177; Conservative 106; Mismatches 214; Indels 44; Gaps 17;

Qy   1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPCIQV 60
          :||::| :|| || :   : : :||: |: :
Db 493 VAPRGKSAAPSKKSK----GCFKEEGVNKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 548

Qy   61 HEDYDCTLNQTNIGNNNKFYIIQLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
         :|: :||| : ::| || |||:||| |:: :| ||| |:| : |:|
Db 549 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRLGTVIGSNKLEQMPSKEE 608

Qy 118 AKKDFFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKP- 176
         | : | | :|||: :: |:||:|: : |||| |:|| | ||| : |:
Db 609 AVEQFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL------TVKPG 659

Qy 177 --CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234
             | |: | :: : |||||||||:||| :||:: :
Db 660 TKSKLPKPVQELVGMIFDVDSMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQPV 719

Qy 235 KNPTGDGQSLEELSSCFYTVTIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
             | ::: ||| | |||: ||::|| :||| :|:|| |||: :|| |:|||||:|
Db 720 SQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLR 778

Qy 295 APGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNL 350
          || :::| ||||||:||: : |:| | :: |||| || ||
Db 779 GGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDL-EV 832

Qy 351 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK 406
          ::: || ||||| :||| ||| ||||||: | | |:||:|||||||:
Db 833 IDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGK 892

Qy 407 GIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIA 466
         |||| :: |||: | | ||:| || |:|||||: ::   :|:|| |
Db 893 GIYFADMVSKSANYCHTSQ--GDPIGLIMLGEVALGNMYELK-HASHISKLPKGKHSVKG 949

Qy 467 RGQTEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRYLLE 525
          | |||| : :| | | ||: : | | :|| || :|: | |:||:||::|| ||||:
Db 950 LGKTTPDPSASITLE----GVEVPLGTGI--PSGVNDTALLYNEYIVYDIAQVNLKYLLK 1003

Qy 526 I 526
         |
Db 1004 L 1004

RESULT 8
S26057
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 18-Jun-1999
A;Accession: S26057; S78453; I52331
R;Thibodeau, J.; Gradwohl, G.; Dumas, C.; Clairoux-Moreau, S.; Brunet, G.; Penning, C.;
Biochem. Cell Biol. 67, 653-660, 1989
A;Title: Cloning of rodent cDNA encoding the poly(ADP-ribose) polymerase catalytic domai
A;Reference number: I52331; MUID:90027702
A;Accession: S26057
A;Molecule type: mRNA
A;Residues: 1-500 <THI>
A;Cross-references: EMBL:X65497; NID:g56849; PIDN:CAA46478.1; PID:g56850
R;Potvin, F.
submitted to the EMBL Data Library, March 1992
A;Reference number: S78453

A;Accession: S78453
A;Molecule type: mRNA
A;Residues: 1-124,'H',126-127,'A',129-238,'D',240-500 <POT>
A;Cross-references: EMBL:X65497
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltrans

Query Match   24.6%;   Score 686; DB 2; Length 500;
Best Local Similarity 36.2%; Pred. No. 9.9e-42;
Matches 173; Conservative 86; Mismatches 179; Indels 40; Gaps 17;

Qy   64 YDCTLNQTNIGNNNKFYIIQLLE-EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKK 120
         : ||| : ::| || |||||| ||| ||:| |||| |:| :| |:|
Db   39 FSATLGLVDIVKGTNSYYKLQLLESDKESRYWIFRSWGRVGTVIGSNKLEQMPSKEDAVE 98

Qy 121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKP---C 177
         :|:: | |||: :: |:||:|: : |||| |:|| | ||| : |:
Db   99 HFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL------AVKPGTKS 149

Qy 178 SLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNP 237
          | |: | :: : |||||||||:||| :||:: :
Db 150 KLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQAVSQG 209

Qy 238 TGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPG 297
         :: ||| | |||: ||::|| :||| :|:|| |||: :|| |:|||||:|
Db 210 SSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNTDSVQAKVEMLDNLLDIEVAYSLLRGGS 268

Qy 298 EEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLRHV 353
          ::| ||||||:||: : |:| | :: |||| || ||
Db 269 DDSSK-----DPIDVNYELKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDL-EVIDI 322

Qy 354 WKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIY 409
         :: || ||||| :||| ||| ||||||: | | |:||:|||||||:
Db 323 FKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIY 382

Qy 410 FASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQ 469
         | :: |||: | | ||:| || |:|||||: ::   :|:|| |
Db 383 FADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSVKGLGK 439

Qy 470 TEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRYLLEI 526
         ||| : :| | | ||: : | | :|| || :|: | |:||:||::|| ||||:
Db 440 TAPDPSASI--TLDG--VEVPLGTGI--PSGVNDTCLLYNEYIVYDIAQVNLKYLLKL 491

RESULT 9
S31735
NAD+ ADP-ribosyltransferase; poly ADP-ribose polymerase
N;Alternate names: poly ADP-ribose polymerase
C;Species: Xenopus laevis (African clawed frog)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
A;Accession: S31735; PN0495
R;Sauller-le Drean, B.M.
submitted to the EMBL Data Library, May 1992
A;Reference number: S31735
A;Accession: PN0495
A;Molecule type: mRNA
A;Accession: S31735
A;Residues: 1-998 <SAU>
A;Cross-references: EMBL:Z12139; NID:g64967; PIDN:CAA78126.1; PID:g1334661
R;Ozawa, Y.; Uchida, K.; Uchida, M.; Ami, Y.; Kushida, S.; Okada, N.; Miwa, M.
Biochem. Biophys. Res. Commun. 193, 119-125, 1993
A;Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)polymera
A;Reference number: PN0494; MUID:93277538
A;Accession: PN0495
A;Molecule type: mRNA
A;Residues: 742-745,'E',747-876 <OZA>
C;Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and diff
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pento

Query Match   24.4%;   Score 681; DB 2; Length 998;
Best Local Similarity 33.6%; Pred. No. 6.1e-41;

Left column

Matches 178; Conservative 95; Mismatches 215; Indels 42; Gaps 18;

```
Qy    10 QTEGSKKQRGT-EEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHEDYDCTL 68
            ::.: .:: .: .|| |            :.:: :: |              :| ||
Db   488 KSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDS--ELEDSCHVLETGG----KIFSATL 540

Qy    69 NQTNIGNNNNKFYIIQLLE--EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKKDFKKK 125
            :| : |:||: |  |:::||    :|||||  |||||  :|::|||||||   ||:| :|
Db   541 GLVDITRGTNSYYKLQLIEHDRSRYWFWFRSWGRVGTVIGSKKLEEMSSKEDAIEHFLNL 600

Qy   126 FWEKTKNKWEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKVDSGPVRTVVKPCSLDPATQ 184
            |||||||||||||||||||:|||| |:|   |||||||||:|||||||||||||     |:
Db   601 YQDKTGNAWHS-PNFTKYPKKFYPLEIDYGQEEDVVKKLSVGAGTKSKLAKP------VQ 653

Qy   185 NLTNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSL 244
            |||||||||||||||||||:||||||||||||:::::|:::|:|||||||||||||||:
Db   654 ELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQAVSESLSEARLL 713

Qy   245 EELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVE 304
            ::.: |||||||||||||||| || |:|::|   :|| |: :.: || |||||||||
Db   714 -DLSNQFYTLIPHDFGMKKFPLLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDGEK-- 770

Qy   305 EVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLK----QTGNSYRCPNLRHVWKVNREG 360
            :::||||||||||||||:||||||||||||:|||      |:||||||||||||||||||
Db   771 ---DPIDVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNAYDLEVL-EIFKIDREG 826

Qy   361 EGDRFQAHSKLGNRRLLWHGTNVAVAAILTSGLRIMPH----SGGRVGKGIYFASENSK 416
            | : :::||||||||||||||||||||:||||||||||    |||||||||||||:|
Db   827 EYQRYKPFKQLHNRQLLWHGSRTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSK 886

Qy   417 SAGYVTTMHCGGHQVGYMFLGEVALGKEHHITDDPSLKSPPPGFDSVIARGQTEPDPAQ 476
            | |:||||||||||||||||||||||||||||:|||:|||||| |||||| ::|| || :
Db   887 SANYCHAM--PGSPIGLILLGEVALGNMHELKAASQITKL-PKGKHSVKGLGRTAPDPSA 943

Qy   477 DIELELDGQPVVVPQGPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI 526
            :|:|||||||||||||||||||||||||||||||||:|||||||||::::
Db   944 TV--QLDG--VDVPLGKGTSA-NISDTSLLYNEFIVYDIAQVNLKYLLKL 988
```

RESULT 10
S42208
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - flesh fly (Sarcophaga peregrina)
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Sarcophaga peregrina
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S42208; S71496
R;Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurata [Sarc...]
Eur J. Biochem. 220, 607-614, 1994
A;Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from Sarc
A;Reference number: S42208; MUID:94170813
A;Accession: S42208
A;Molecule type: mRNA
A;Residues: 1-996 <MAS>
A;Cross-references: EMBL:D16482; NID:g473742; PIDN:BAA03943.1; PID:g538248
A;Accession: S71496
A;Molecule type: protein
A;Residues: 170-188;721-736;813-819;879-885 <MAX>
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger
F;1-369/Domain: DNA binding #status predicted <DNA>
F;370-507/Domain: auto-modification #status predicted <AMO>
F;508-996/Domain: NAD binding #status predicted <NAD>

Query Match 23.4%; Score 652; DB 2; Length 996;
Best Local Similarity 33.3%; Pred. No. 7.6e-39;
Matches 176; Conservative 100; Mismatches 205; Indels 48; Gaps 18;

```
Qy    22 EEEDSFRSTAEALRAAPADNRV----IRVDPSCPF------IRVDPSCPF-----SRNPGIQVHEDYDCTL 68
            || || || :  : :: :::  :        ||| |        ::: ||
Db   489 EESKSSKSKSIYTKSVPKSMTLKIKDGLAVDPDSGLEDVAHVVYSRN----KEKYNVVL 543

Qy    69 NQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKKK 125
```

Right column

```
Db   544 GITDIQKNKNSFYKLQLLESDMKNRFWVFRSWGRIGTTIGGNKLDNFSNLVDAIVQFKEL 603

Qy   126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLDPATQN 185
            :|| :  |:.:  | ::     :||:::|||||||||||||||||||||||||||||
Db   604 YLEKSGNHFENRENFVKVAGRMYPIDIDYAEDS----KIDLSAEHDIKSKLPL--SVQD 656

Qy   186 LITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSLE 245
            :|  |||||||||||||||||||||||||||||||:::||:::|:||||| |: :|
Db   657 IIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLTEIYELIQG-GGTNAKFI 715

Qy   246 ELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADELAQTLQAAPGEEEEKVEE 305
            :::|||||||||| ||| : :  :    :|:   |:::|:|::|:|: |::  |||||
Db   716 DATNRFYTLIPHNFGTQSPPLLDTTEQVEQLRQMLDSLIEIECAYSLLQT---EDSKAD- 771

Qy   306 VPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR-HVWKVNREGEG 362
            :: : :::.: :: ||:      |||| ||:     ::|||:   |   .::||||||:|
Db   772 -INPIDKHYEQLKTKLEPLDKNSEEYILLQKYVKNTHAETHKLYDLEVVDIFKVARQGEA 830

Qy   363 DRFQAHSKLGNRRLLWHGTNVAVAAILTSGLRIMPH----SGGRVGKGIYFASENSKSA 418
            |||||||||||||||||||||||||||||||||||    |||||||||||||:|:|||
Db   831 RRYKPFKKLHNRRLLWHGSRLTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADMVSKSA 890

Qy   419 GYVTTMHCGGHQVGYMFLGEVALGKEHHITDDPSLKSPPPGFDSVIARGQTEPDPAQDI 478
            ||| ||||||||||||||||||||||||||:|||:|||||| |||||:| :|| ||||
Db   891 NYCCTSH--HNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFGRGRTMPNPSESI 947

Qy   479 ELELDGQPVVVPQGPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIH 527
            :|:|||||||||||||||||||||||||||||| :|||||||||||:::
Db   948 IRE-DG--VEIPLGKPITNDSLK-SSLLYNEFIIYDIAQVNIQYMLRMN 992
```

RESULT 11
T51353
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) [imported] - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C;Accession: T51353
R;Doucet-chabeaud, G.; Kazmaier, M.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z25379
A;Accession: T51353
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-983 <DOU>
A;Cross-references: EMBL:AJ131705; PIDN:CAA10482.1
A;Experimental source: cultivar landsberg erecta
C;Genetics:
A;Gene: parp-1
C;Function:
A;Description: ADP-ribose polymer synthesis
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match 22.4%; Score 624; DB 2; Length 983;
Best Local Similarity 33.0%; Pred. No. 7.8e-37;
Matches 179; Conservative 87; Mismatches 209; Indels 68; Gaps 21;

```
Qy    15 KKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH----ED----YDC 66
            ::  || ::   :|:::|: |  |: :         |||:::|  |     ||    |:
Db   472 KKQRKLPFDKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT 524

Qy    67 TLNQTNIGNNNNKFYIIQLLEE--GSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDF 122
            ||| :|:|||||||||||||||    |:||||||| |||  :|::|||:|:: |::::
Db   525 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGGNKVEEMS-KSDAVHEF 583

Qy   123 KKKFWEKTKN--KWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVKPCSL 179
            :| | ||::    ::|| | :|: :  : :.:: |:||||||:||||||||||||||:
Db   584 KRLFLEKTGNIWESWEQKTNFQKQPGKFLPLDIDYGVNKQVA----KKEPFQT----SSNL 636

Qy   180 DPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL----EALEEAMK 235
```

```
Qy  637 APSLIELMKLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQRLLTESDP 696
            |:  |:  |   | :::|         |||||||:| ||||| ||:|||||||||
Db  236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295

Db  697 QPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIEIASRI--- 749
            ||   |   : | |:|:||    |||||||||||||||||:| |||||:|||||:
Qy  296 PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------- 348

Db  750 VGFDVDSTES----LDDKYKKLHCDISPLPHDSEDYRLIEKYL----NTTHAPTHTEWSL 801
            ||      |   || : |::: |  || :  |||||:|||||    ||||||||||||
Qy  349 NLRHVWKVNREGEGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRI---MPHSGGR 403

Db  802 ELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAPATGYM 861
            :|||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Qy  404 VGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDS 463

Db  862 FGKGIYFADLVSKSAQYCYT--CKKNPVGLMLLSEVALGEIHELT-KAKYMDKPPRGKHS 918
            ||||||||||||||||||||  ||||||||||||||||||||||| |||||||||||||
Qy  464 VIARGQTEPDPAQDIELELDGQPVVVPQGPPFSFKSSSFSQSEYLIYKESQCRLRYL 523

Db  919 TKGLGKKVP---QDSEFAKWRGDVTVPCGKPVS-SKVKASELMYNEYIVYDTAQVKLQFL 974
            ||||||||    |||||||||||||||||||||  ||||||||||||||||||||||||
Qy  524 LEI 526
            |:
Db  975 LKV 977

RESULT 12
C84719
probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: C84719
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Shen, J.A.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: C84719
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1009 <STO>
A:Cross-references: GB:AE002093; NID:g4432827; PIDN:AAD20677.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g31320
A;Map position: 2
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match       22.4%;    Score 624;    DB 2;    Length 1009;
Best Local Similarity   33.0%;    Pred. No. 8.1e-37;
Matches  179;   Conservative  87;   Mismatches 209;   Indels 68;   Gaps 21;

Qy   15 KKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH---ED----YDC 66
            |||  |  :    | ||:  ||      | | :   :|: :||||:      :||
Db  498 KKQRKLPFDVKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT 550

Qy   67 TLNQTNIGNNNKFYIIQLLEE--GSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDF 122
            ||   :  ::| ||||||   | |  ||| |||||  | ||||  :|||||||||||||
Db  551 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGGNKVEEMS-KSDAVHEF 609

Qy  123 KKKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSL 179
            |: | ||| |    |  : ::    :  | :  :  :  |  :|  : | :|||| |
Db  610 KRLFLEKTGNTWESWEQKTNFQKPQGKFLPLDIDYGVNKQVA---KKEPFQT---SSNL 662

Qy  180 DPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL----EALEEAMK 235
            :|||  |:||:|||||||||||| ||:|:|||||||:|:|| |||||     ||||||||
Db  663 APSLIELMKLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQRLLTESDP 722

Qy  236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
            ||   |  | ||: :  || : |  ||
```

```
Db  723 QPTMKESLLVDASNRFFTMIP----SIHPHIIRDEDDFKSKVKMLEALQDIEIASRI--- 775
            ||   |   : | |:|:||     |||||||||||||||||:| |||||:|||||:
Qy  296 PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------- 348

Db  776 VGFDVDSTES----LDDKYKKLHCDISPLPHDSEDYRLIEKYL----NTTHAPTHTEWSL 827
            ||      |   || : |::: |  || :  |||||:|||||    ||||||||||||
Qy  349 NLRHVWKVNREGEGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRI---MPHSGGR 403

Db  828 ELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAPATGYM 887
            :|||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Qy  404 VGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDS 463

Db  888 FGKGIYFADLVSKSAQYCYT--CKKNPVGLMLLSEVALGEIHELT-KAKYMDKPPRGKHS 944
            ||||||||||||||||||||  ||||||||||||||||||||||| |||||||||||||
Qy  464 VIARGQTEPDPAQDIELELDGQPVVVPQGPPFSFKSSSFSQSEYLIYKESQCRLRYL 523

Db  945 TKGLGKKVP---QDSEFAKWRGDVTVPCGKPVS-SKVKASELMYNEYIVYDTAQVKLQFL 1000
            ||||||||    |||||||||||||||||||||  ||||||||||||||||||||||||
Qy  524 LEI 526
            |:
Db 1001 LKV 1003

RESULT 13
A47474
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: A47474
R:Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.
Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993
A:Title: Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase: leucine zip
A:Reference number: A47474; MUID:93234521
A:Accession: A47474
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1994 <UCH>
A:Cross-references: GB:D13806; GB:D13807; GB:D13808; NID:g303545; PIDN:BAA02964.1; PI
A:Note: sequence extracted from NCBI backbone (NCBIN:129703, NCBIP:129704)
C;Genetics:
A;Gene: FlyBase:Parp
A;Cross-references: FlyBase:FBgn0010247
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match       22.3%;    Score 621.5;    DB 1;    Length 994;
Best Local Similarity   31.6%;    Pred. 1.2e-36;
Matches  167;   Conservative 102;   Mismatches 221;   Indels 39;   Gaps 17;

Qy   16 KQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIO----VHED----YDCT 67
            | :   |      || :   ||  :  :| :     ||: :||:||        |:|
Db  481 KSRIPKETTKSLNSNSIYTKSMPV-SRTFKVKDGLAVDPDSGLEDIAHYVDSNNKYSVV 539

Qy   68 LNQTNIGNNNKFYIIQLL--EEGSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKK 124
            |  | |  ||| ||:|||| :  | ||||| ||:|| :::|||:|:|||:|||: :||:
Db  540 LGLTDIQRNKNSYYKVQLLKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKE 599

Qy  125 KFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLDPATQ 184
             |:|||| | |:|:|:|     | :|:||  | :||  |  ::::|   |  |:       |
Db  600 IYADKTGNEYEQRDNFVKRTGRMYPIEIQ--YDDDQKLVKHESHFFTS----KLEISVQ 652

Qy  185 NLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEEEAMKNPTGDGQSL 244
            ||| :|| ||| | |||| |  ::|||||||: :|  :   ::  : :|
Db  653 NLIKLIFDIDSMNKTLMEFHIDMDKMPLGKLSAHIQSAYRVKEIYNVLECGSNTAK-L 711

Qy  245 EELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVE 304
            : |: : :||||||||  :||: ::    || :||::| :|:| :  :|
Db  712 IDATNRFYTLIPHNFGVQLPTLIETHQQIEDLRQMLDSLAEIEVAYSI----IKSEDVS 766

Qy  305 EVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR-HVHKVNREGE 361
            :  : ||::|||:   ::|| |:|::|  | : :|  | ||||| |||  |:|||:|||
Db  767 DACNPLDNHYAQIKTQLVALDKNSEEFSILSQVVKNTHASTHKSYDLKIVDVFVVSRQGE 826
```

```
Qy    362  GDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRI----MPHSGRVGKIYFASENSKS  417
           ||  :  :::|||||||:     ||  ||    |     ||:||| |||||||||
Db    827  ARRFKPFKKLHNRKLLWHGSRLTNFVGILSHGLRLRIAPPEAPPTGYMFGKGIYFADMVSKS  886

Qy    418  AGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPGFDSVIARGQTEPDPAQD  477
           | |   |||      |  ||||  ||    |   |  :||||||||| |||||||| |||
Db    887  ANYCCTSQ--QNSTGLMLLSEVALGDMMECT-SAKYINKLSNNKHSCFGRGRTMPDPTKS  943

Qy    478  IELELDGQPVVPVQPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI  526
           :  ||   :  |:|:   :  || ::|||: |:: ||
Db    944  Y-IRSDG--VEIPYGETITDEHLK-SSLLYNEYIVYDVAQVNIQYLFRM  988
```

RESULT 14
T03657
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) 2 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
C;Accession: T03657
R;Bablychuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Monta
submitted to the EMBL Data Library, November 1997
A;Description: Higher plants possess two poly(ADP-ribose) polymerases.
A;Reference number: 214992
A;Accession: T03657
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-969 <BAB>
A;Cross-references: EMBL:AJ222589; PIDN:CAA10889.1
C;Genetics:
A;Gene: PARP2
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

```
Query Match        20.9%;  Score 584;  DB 2;  Length 969;
Best Local Similarity   30.9%;  Pred. No. 5.9e-34;
Matches 164;  Conservative  94;  Mismatches 201;  Indels  72;  Gaps 20;

Qy     33  ALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED----YDCTLNQTNIGNNNNKFYIIQ  84
           ||       ::|    :  : :|  |  :: | : ||:::  :::::|
Db    468  ALESSKGSTVTVKVKGRSAVHESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQ  527

Qy     85  LLE--EGSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDFKKKFWEKTKNKWEE---R  137
           |||     |: :  :  |  || ::    |  | || |:::  | | ||:| |:   |
Db    528  IIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEKTGNSWEAWECK  586

Qy    138  DRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVK--PCSLDPATQNLITNIFSKEM  195
           |     :: ::   ||  :      ::| |  |   : :   : ::|:: | ::::
Db    587  TNFRKQPGRFYPLDVD--------YGVKKAPKKDISEMKSSLAPQLLELMKMLFNVET  637

Qy    196  FKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSLEE----LSSC  250
           |   | |  || | |:  :||||:  | | |:| |   ||   |  ::|:    ||:
Db    638  YRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAASNR  697

Qy    251  FYTVIPHNFGRSRPPINSPDVLVLADIELAQTLQAAPGEEEEKVEEVPHPL  310
           |:|:|   |||::|  ||::||| |||:||||  |   ::|:
Db    698  FFTLIP----SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDES------L  746

Qy    311  DRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------NLRHVWKVNREGEGD  363
           :|:||:|||||| |||||||:|:  ||| |   ::|
Db    747  DDKYMKLHCDITPLAHDSEDYKLIEQYLLNT----HAPTHKDWSLELEEVFSLDRDGELN  802

Qy    364  RFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKIYFASENSKSA  418
           |||   ||||||||||||    | || |  |||::      :||||||||||||||| |
Db    803  KYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVSKSA  862

Qy    419  GYVTTMHC--GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPGFDSVIARGQTEPDPA  475
           ||:  |||    | | ||| ||||| ||     ||||||||||||| |||||||| ||
Db    863  QY-----CYVDRNNPVGLMLLSEVALGDMYELK-KATSMDKPPRGKHSTKGLGKTVPLES  916

Qy    476  QDIELELDGQPVVPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI  526
             ::  |||||  | | |||:  |||:|||||||||||||||||::
```

```
Db    917  EFVKWRDD--VVVPCGKPVP-SSIRSSELMYNEYIVYNTSQVKMQFLLKV  963
```

RESULT 15
T18600
hypothetical protein AC8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18600
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: 218996
A;Accession: T18600
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-727 <WIL>
A;Cross-references: EMBL:Z83097; PIDN:CAB05448.1; GSPDB:GN00023; CESP:AC8.1
A;Experimental source: clone AC8
C;Genetics:
A;Gene: CESP:AC8.1
A;Map position: 5
A;Introns: 31/3; 271/3; 312/3; 348/1; 466/3; 545/3; 617/3; 702/2

```
Query Match        17.2%;  Score 481;  DB 2;  Length 727;
Best Local Similarity   30.0%;  Pred. No. 1.1e-26;
Matches 133;  Conservative  68;  Mismatches 165;  Indels 78;  Gaps 13;

Qy    126  FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLDPATQN  185
           |  | ||:|||   :  |||||||| :||  ||| |   ||| ||| | ||      :|
Db    315  FHEKTKNDWIYRKHFRKMPGMFSYVET----DYSEFVGTNNGHKKKITPGSKITPGSKT  369

Qy    186  L-----ITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNP  237
           |       ::| ::  :| : ||:|  |||||||| |||:||| ||:||||  |:|
Db    370  LLPKSVKEVMSIFDVENMKSALKSFEIDVNKMPLGRLSHNQINLAFEVLNDISDLLVKL  429

Qy    238  TGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPG  297
           ||||| ||| ||:|:|:|  ||||:||||||||||  : || :|||||||||| :|
Db    430  PIDASKILDFSNKFYTIPHNFGMRVPEPIDSFHKIKEKNMLNALLDIKFAYD-QISGG  488

Qy    298  EEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCP-NLRHVWK  355
           :  ||:|  |  |||| |||||| | |||||||| ||||| ||    :| | ||||:|
Db    489  DVPASTSLSIDPVDINYRKLKCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVELIDILK  548

Qy    356  VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFA  411
           |||  |   ||  |||||||||||| || | || ||||||:      |||||||| |||
Db    549  VNRDNESSKFKRH-IGNRRLLWHGSGKMNFAGILGQGLRIAPPEAPVSGYMFGKGVYFA  606

Qy    412  SENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFGFDSVIARGQT  470
           ||| |||| ||||| |||||||||| |||| |||||| | ||||  | |||||||:|||| |
Db    607  DMFSKSFFY--CRANAKEFAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSV------  657

Qy    471  EPDPAQDIELELDGQPVVVPQGPPVQCP--SFKSS--------SFSQ-  507
            |||| || ||||||| | || |||| |  ||||         |||
Db    658  -------------QGVGRQCPREIGSYKNPDGYTVPLGLTYMQLGKQNVDYH  697

Qy    508  --SEYLIYKESQCRLRYLLEIHL  528
             ::|||| :::|||||:|: :
Db    698  LLYNEFIVYDQIQLKYLVRVKM  721
```

Search completed: August 29, 2002, 07:59:41
Job time: 288 sec

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

```
          GenCore version 4.5
    Copyright (c) 1993 - 2000 Compugen Ltd.
```

Run on: August 29, 2002, 08:02:24 ; Search time 36.19 Seconds
                                           (without alignments)
                                  564.905 Million cell updates/sec

Title: US-09-701-586B-10
Perfect score: 2789
Sequence: 1 MAPKRKASVQTEGSKKQRQG.........EYLIYKESQCRLRYLLEIHL 528

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2241.5	80.4	533	1	PPO3_HUMAN	Q9y6f1 homo sapien
2	718	25.7	1012	1	PPOL_CRIGR	Q9r152 cricetulus
3	712.5	25.5	559	1	PPO2_MOUSE	O88554 mus musculu
4	707.5	25.4	1013	1	PPOL_HUMAN	P09874 homo sapien
5	704.5	25.3	1015	1	PPOL_BOVIN	P18493 bos taurus
6	699	25.1	583	1	PPO2_HUMAN	Q9ugn5 homo sapien
7	698	25.0	1013	1	PPOL_RAT	P27008 rattus norv
8	697.5	25.0	637	1	PPOL_ARATH	Q11207 arabidopsis
9	694	24.9	1011	1	PPOL_CHICK	P26446 gallus gall
10	688	24.7	1012	1	PPOL_MOUSE	P11103 mus musculu
11	681	24.4	998	1	PPOL_XENLA	P31669 xenopus lae
12	652	23.4	996	1	PPOL_SARPE	Q11208 sarcophaga
13	621.5	22.3	994	1	PPOL_DROME	P35875 drosophila
14	389	13.9	538	1	YQN1_CAEEL	O09525 caenorhabdi
15	276.5	9.9	1724	1	PPOV_HUMAN	Q9ukk3 homo sapien
16	193	6.9	135	1	PPOL_ONCMA	Q08824 oncorhynchu
17	113.5	4.1	2035	1	EVPL_MOUSE	Q9d952 mus musculu
18	107.5	3.9	638	1	TOXA_PSEAE	P11439 pseudomonas
19	106.5	3.8	646	1	YE14_SCHPO	O13869 schizosacch
20	105.5	3.8	3075	1	LMA1_HUMAN	P25391 homo sapien
21	103	3.7	379	1	YO11_MOUSE	P11260 mus musculu
22	103	3.7	1031	1	KINH_STRPU	P35978 strongyloce
23	102.5	3.7	612	1	EXO2_BPT5	P11109 bacteriopha
24	101	3.6	1574	1	RPOC_AQUAE	O67763 aquifex aeo
25	100.5	3.6	629	1	PAB2_ARATH	P42731 arabidopsis
26	100.5	3.6	830	1	YBMA_SCHPO	Q10332 schizosacch
27	100.5	3.6	890	1	RB6K_HUMAN	O95235 homo sapien
28	100.5	3.6	1395	1	SP41_YEAST	P38904 saccharomyc
29	100.5	3.6	3210	1	CENF_HUMAN	P49454 homo sapien
30	100	3.6	2025	1	TTC3_HUMAN	P53804 homo sapien
31	99	3.5	583	1	T2F1_FLAOK	P14870 flavobacter
32	98	3.5	2453	1	NCRL_MOUSE	Q60974 mus musculu
33	97.5	3.5	548	1	SYPB_AERPE	Q9y913 aeropyrum p
34	97	3.5	727	1	MFP1_ARATH	Q91w85 arabidopsis
35	96.5	3.5	495	1	NUSA_ECOLI	P03003 escherichia
36	96.5	3.5	497	1	BIR4_HUMAN	P98170 homo sapien
37	96.5	3.5	583	1	LAM1_XENLA	P09010 xenopus lae
38	96.5	3.5	808	1	PML_MOUSE	Q60953 mus musculu
39	96.5	3.5	853	1	PBPA_HAEIN	P31776 h penicilli
40	96.5	3.5	1396	1	RPOC_CHLMU	Q9pk79 chlamydia m
41	96.5	3.5	3587	1	SRF2_BACSU	Q04747 bacillus su
42	96	3.4	605	1	VCLA_GOSHI	P09799 gossypium h
43	96	3.4	963	1	KINH_HUMAN	P33176 homo sapien
44	96	3.4	1875	1	MLP1_YEAST	Q02455 saccharomyc
45	95.5	3.4	725	1	ADDB_MOUSE	Q9qyb8 mus musculu

ALIGNMENTS

RESULT   1
PPO3_HUMAN
ID   PPO3_HUMAN   STANDARD;   PRT;   533 AA.
AC   Q9Y6F1; Q9UG81;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+) ADP-
DE   ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (pADPRT-3)
DE   (hPARP-3).
GN   ADPRTL3 OR PARP3 OR ADPRT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=99263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [2]
RP   SEQUENCE OF 75-533 FROM N.A.
RC   TISSUE=Kidney;
RA   Ansorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl (N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels are in
CC   the kidney, skeletal muscle, liver, heart and spleen; also
CC   detected in pancreas, lung, placenta, brain, leukocytes, colon,
CC   small intestine, ovary, testis, prostate and thymus.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF083068; AAD29855.1; -.
DR   EMBL; AL050034; CAB43246.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein.
KW   ADP-ribosylation.
FT   DOMAIN   14   20   NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   CONFLICT   80   80   K -> N (IN REF. 2).

```
FT   CONFLICT      171            171         A -> G (IN REF. 2).
FT   CONFLICT      411            411         K -> E (IN REF. 2).
SQ   SEQUENCE   533 AA;  60117 MW;  7C0AB89E64D1B9FD CRC64;

Query Match             80.4%;  Score 2241.5;  DB 1;  Length 533;
Best Local Similarity   80.5%;  Pred. No. 4.4e-157;
Matches  430;  Conservative  37;  Mismatches  60;  Indels  7;  Gaps  4;

Qy     1  MAPKRKASVQTEG--SKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI  58
          |||| | |||||    || |||||||||||:| |||||||:|| | |||
Db     1  MAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT  60

Qy    59  QVHEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117
          |:||| | || ||| || |||||||||:::::  :||  |||||||||:|||:|||||
Db    61  QVYEDYNCTLNQTNIENNNKKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120

Qy   118  AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTV---V 174
          ||||:||| |||||||||||| | :|| | :|||||  |:|||||||||||| ||
Db   121  AKKDFEKKFREKTKNNWAEERDHFVSHPGKYTLIEVQADEAQEAVVKVDRAPVRTVTKRV 180

Qy   175  KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEBAM 234
          :|||||||| ||||||||||||| |  |||||||||||:||||||||||||||||||:
Db   181  QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240

Qy   235  KNPIGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
          |||  |||||||:|| || ||||||  |:||||||||| ||||:|||||||||| |||:
Db   241  KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 300

Qy   295  APGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGSEYKAIQTYLKQTGNSYRCPNLRHVW  354
          |  |:|| ||||||||||||||||||||||:|   ||:||| |:|:  |||:| |
Db   301  V-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIW  359

Qy   355  KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 414
          |||  :||||||||||||| |||||||  ||||||||||||||||||||||:||||||||
Db   360  KVNQGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 419

Qy   415  SKSAGYVTTMHCGGHQVGMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
          ||||||| | || |:|||||||||||| ||:| |||:|||||:|||:||||| |||||
Db   420  SKSAGYVIGMKCGAHHVGMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP 479

Qy   475  AQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
          | |||||||||  ||||| || ||  |||||||:||||||:||||||||: :||
Db   480  TQDTELELDGQQVVVPQGPVPCPEFSSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
```

RESULT   2
PPOL_CRIGR

```
ID   PPOL_CRIGR     STANDARD;      PRT;   1012 AA.
AC   Q9R152;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21276334; PubMed=11382339;
RA   Ganesh A., Phillips E., Thacker J., Meuth M.;
RT   "Suppression of the radiation-sensitive phenotype of hamster irs1 and
RT   irs2 strains selected for resistance to 3-aminobenzamide.";
RL   Int. J. Radiat. Biol. 77:609-616(2001).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use  by  non-profit institutions as  long  as its content is in no  way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF168781; AAD45817.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0
FT   DNA_BIND      0    371       BY SIMILARITY.
FT   DOMAIN      372    522       AUTOMODIFICATION DOMAIN (BY SIMILARITY).
FT   DOMAIN      386    475       BRCT.
FT   DOMAIN      523   1012       NAD-BINDING (BY SIMILARITY).
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       PARP-TYPE.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                (POTENTIAL).
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                (POTENTIAL).
FT   MOD_RES     406    406       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     412    412       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     443    443       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     455    455       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     470    470       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     483    483       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     487    487       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     490    490       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     511    511       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N]  (POTENTIAL).
FT   MOD_RES     518    518       ADP-RIBOSYL[N]  (POTENTIAL).
SQ   SEQUENCE   1012 AA;  112400 MW;  8D617C4DBF0CB0F7 CRC64;

Query Match             25.7%;  Score 718;  DB 1;  Length 1012;
Best Local Similarity   34.3%;  Pred. No. 7e-45;
Matches  185;  Conservative  98;  Mismatches  213;  Indels  44;  Gaps  17;

Qy     2  APKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH  61
          ||| |::    :: ||          :  ||        :            :
```

```
Db  493 APKGKSAAPSKKSK----GLYKEEGVNKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKGG 548
         :  |     |          | ||   |:: |||:  | :|| |  : ||  :|:|:|
Qy   62 EDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLEDA 118

Db  549 KVFSATIGLVDIVKGTNSYYKIQLLEDDKESRYWIFRSWGRVGTVIGSNKLEQMPSKEDA 608
         :     | :|    :    |  |  | |      ||  :||   |
Qy  119 KKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKP-- 176

Db  609 VEHFMKILYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL------TVKPGT 659
          |           |    | |  ||  |  |||  |     ||  :        ||
Qy  177 -CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 235

Db  660 KSKLPKAVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQAVS 719
          | |       :       : : |   : :|  ||||||: |:|| | |::|||:|| |
Qy  236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295

Db  720 QGSSDSQIL-DLSNRFYTLIPHDFGWKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLRG 778
              | :   :|||::|  ||| || ::||  |:||:|||  |:|:|| ||||| | | |
Qy  296 PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGSEYKAIQTYLKQ---TGNSYRCPNLR 351

Db  779 GSDDSSK----DPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDL-EVM 832
              | :     | ||:| |||:|||| ||||||||::|:||:|:| |::|:|| | |:|
Qy  352 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVRAILTSGLRIMPH----SGGRVGKG 407

Db  833 DIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKG 892
           :|  |:  : ::  | |:|:||||||||: . : |||| ||||   :  |||||| ||
Qy  408 IYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIAR 467

Db  893 IYFADMVSKSANYCHTSQ--GDPIGLLLGEVALGNMYELK-HASHISKLPKGKHSVKGL 949
          |||| :| |||  || :   ||| | | |||||: ||:||  :|| ||||||||| |||
Qy  468 GQTEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRYLLEI 526

Db  950 GKTTPDPSASITLE----GVEVPLGTGI--PSGVNDTCLLYNEYIVVDIAQVNLKYLLKL 1003
         | |.||||:|:||     |||||| |    ||||||.|||:||||||||||||||||:|
```

```
RESULT  3
PPO2_MOUSE
ID   PPO2_MOUSE    STANDARD;    PRT;    559 AA.
AC   O88554; Q99N29;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE   (mPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Embryo;
RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129Sv;
RX   MEDLINE=21179160; PubMed=11133988;
RA   Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,
RA   Niedergang C.P.;
RT   "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2
RT   (PARP-2) gene to the gene for RNase P RNA.";
RL   J. Biol. Chem. 276:11092-11099(2001).
RN   [3]
RP   SEQUENCE OF 9-559 FROM N.A.
RC   STRAIN=129/Sv X C57BL/6;
RX   MEDLINE=99268466; PubMed=10338144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RL   elegans.";
RL   FEBS Lett. 449:259-263(1999).
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC       SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in
CC       testis followed by ovary.
CC   -!- INDUCTION: By high levels of DNA-damaging agents.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
```

```
DR   EMBL; AJ007780; CAA07679.1; -.
DR   EMBL; AF191547; AAK13253.1; -.
DR   EMBL; AF072521; AAC25415.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1341112; Adprt2.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND      1     65       POTENTIAL.
FT   DOMAIN       66    559       NAD-BINDING (BY SIMILARITY).
FT   DOMAIN        3      9       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                (POTENTIAL).
FT   DOMAIN       33     39       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                (POTENTIAL).
FT   CONFLICT     82     82       L -> V (IN REF. 2).
FT   CONFLICT    177    177       V -> I (IN REF. 2).
FT   CONFLICT    486    486       R -> Q (IN REF. 2).
SQ   SEQUENCE   559 AA;  63396 MW;  E0AEDAEE412C1445 CRC64;

Query Match             25.5%;       Score 712.5;    DB 1;   Length 559;
Best Local Similarity   35.2%;       Pred. No. 7.8e-45;
Matches  204;  Conservative  87;  Mismatches 207;  Indels 81;  Gaps  25;

Qy    1 MAPKRKAS------VQTEGSKKQRQGTE--EEDS--FRSTAEALRAAP----------ADN 41
         |||:|:        | ||::    |||| | ||                         |||
Db    1 MAPRRQRSGSGGRRVLNE-AKKVDNGNKATEDDSPPGKKMRTCQRKGPMAGGKDADRTKDN 59

Qy   42 R---------VIRVDPSCPFSRNPG--IQVHEDYDCTLNQTNIGNNNNKFYIIQLLEE 88
         |         |||          |||| | |||||::|||::|||:|||||||||||
Db   60 RDSVKTLLLKGKAPVDPECAAKLGKAHVYCEGDDVYDVMLNQTNLQFNNNKYYILQLLED 119

Qy   89 GSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFKKKFWEKTKNKWEERDRFVAQ 143
         | |: :| |||||| || |||||:: || :: ||::||  | :|| |:|||:||||||
Db  120 DAQRNFSVWMRWGRVGKTGQHSL--VTCSGDLNKAKEIFQKKFLDKTKNWEDRENFEKV 177

Qy  144 PNKYTLIEVQGEAESQ-EAVVKVDSGPVRTVVKPCS-LDPATQNLITNIPSKEMFKNAMT 201
         ||| ::||: :: | |  :|||||| |||| |   |::|  |  | ||:|  | || | |
Db  178 PGKYDMLQMDYAASTQDESKTKEE-----TLKPESQLDLRVQELLKLICNVQTMEEMMI 232

Qy  202 LMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD-GQSLEELSSCFYTVIPHNFG 260
         |:|||||||||||||:|||| ||||||||||| |:| | | ||||||||||:|||||| |
Db  233 EMKYDTKRAPLGKLTVAQIKAGYQSLKKIEDCIR--AGQHGRALVEACNEFYTRIPHDFG 290

Qy  261 RSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQ 320
         :: ||:|||||  | | |:|| ||:||  | ::| |  | : | :|:|:|||| | |
Db  291 LSIPPVIRTEKELSDKVKLLEALGDIEIALKLVKS---ERQGLE---HPLDQHYRNLHCA 344
```

7

```
Qy  321 LQLLDSGESEYKAIQTYLKQTGNSYRCP-------NLRHVWKVNREGEGDRFQAHSKLGN 373
            |:|| :|:|  |:|:  |   |  :  |:| :|||::|:   ||  |
Db  345 LRPLDHESNEFKVISQYLQST----HAPTHKDYTMTLLDVFEVEKEGEKEAFR--EDLPN 398

Qy  374 RRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSKSAGYVTMHCGGH 429
            |||||| :  :   ||:||:|: :    :  |||||||||:    |: |   |  |
Db  399 RMLLWHGSRLSNWVGILSHGLRVAPPEAPITGYMFGKGIYFADMSSKSANYCFASRL--K 456

Qy  430 QVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPVVV 489
            |  | ||||||  : :  : |    |:    |  :  :  |   | : |  |
Db  457 NTGLLLLSEVALGQCNELLEANPKAQGLLRGKHSTKGMGKMAPSPAHFI--TLNGS--TV 512

Qy  490 PQGPPVQCPSFKSSSF--SQSEYLIYKESQCRLRYLLEI 526
            ||    ::    :   |::::|| :|  :||||:|
Db  513 PLGPASDTGILNPEGYTLNVNEFIYVSPNQVRMRYLLKI 551
```

RESULT 4

PPOL_HUMAN

```
ID   PPOL_HUMAN      STANDARD;      PRT;   1013 AA.
AC   P09874;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT OR PPOL OR PARP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90091744; PubMed=2513174;
RA   Auer B., Nagl U., Herzog H., Schneider R., Schweiger M.;
RT   "Human nuclear NAD+ ADP-ribosyltransferase(polymerizing):
RT   organization of the gene.";
RL   DNA 8:575-580(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88076933; PubMed=3120710;
RA   Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S.,
RA   Suzuki H., Nyunoya H., Miwa M., Sugimura T.;
RT   "Nucleotide sequence of a full-length cDNA for human fibroblast
RT   poly(ADP-ribose) polymerase.";
RL   Biochem. Biophys. Res. Commun. 148:617-622(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88058958; PubMed=2824474;
RA   Kurosaki T., Ushiro H., Mitsuuchi Y., Suzuki S., Matsuda M.,
RA   Matsuda T., Katunuma N., Kangawa K., Matsuo H., Hirose T.,
RA   Inayama S., Shizuta Y.;
RT   "Primary structure of human poly(ADP-ribose) synthetase as deduced
RT   from cDNA sequence.";
RL   J. Biol. Chem. 262:15990-15997(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88068596; PubMed=2891139;
RA   Cherney B.W., McBride O.W., Chen D., Alkhatib H., Bhatia K.,
RA   Hensley P., Smulson M.E.;
RT   "cDNA sequence, protein structure, and chromosomal location of the
RT   human gene for poly(ADP-ribose) polymerase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:8370-8374(1987).
RN   [5]
RP   SEQUENCE OF 440-1013 FROM N.A.
RX   MEDLINE=87298455; PubMed=3113420;
RA   Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA   Miwa M.;
RT   "Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and
RT   expression of its gene during HL-60 cell differentiation.";
RL   Biochem. Biophys. Res. Commun. 146:403-409(1987).
RN   [6]
RP   ERRATUM.
RA   Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA   Miwa M.;
RL   Biochem. Biophys. Res. Commun. 148:1549-1550(1987).
RN   [7]
RP   SEQUENCE OF 1-94 FROM N.A.
RX   MEDLINE=91099327; PubMed=2125269;
RA   Yokoyama Y., Kawamoto T., Mitsuuchi Y., Kurosaki T., Toda K.,
RA   Ushiro H., Terashima M., Sumimoto H., Kuribayashi I., Yamamoto Y.,
RA   Maeda T., Ikeda H., Sagara Y., Shizuta Y.;
RT   "Human poly(ADP-ribose) polymerase gene. Cloning of the promoter
RT   region.";
RL   Eur. J. Biochem. 194:521-526(1990).
RN   [8]
RP   SEQUENCE OF 1-39 FROM N.A.
RX   MEDLINE=90211250; PubMed=2108670;
RA   Ogura T., Nyunoya H., Takahashi-Masutani M., Miwa M., Sugimura T.,
RA   Esumi H.;
RT   "Characterization of a putative promoter region of the human
RT   poly(ADP-ribose) polymerase gene: structural similarity to that of
RT   the DNA polymerase beta gene.";
RL   Biochem. Biophys. Res. Commun. 167:701-710(1990).
RN   [9]
RP   SEQUENCE OF 1-39 FROM N.A.
RA   Herzog H., Schneider R., Hirsch-Kauffmann M., Schnitzer D.,
RA   Schweiger M.;
RL   Submitted (JUL-1991) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=90222155; PubMed=2109322;
RA   Gradwohl G., Menissier de Murcia J., Molinete M., Simonin F.,
RA   Koken M.H.M., Hoeijmakers J.H.J., de Murcia G.M.;
RT   "The second zinc-finger domain of poly(ADP-ribose) polymerase
RT   determines specificity for single-stranded breaks in DNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:2990-2994(1990).
RN   [11]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=91072398; PubMed=2123876;
RA   Ikelima M., Noguchi S., Yamashita R., Ogura T., Sugimura T.,
RA   Gill D.M., Miwa M.;
RT   "The zinc fingers of human poly(ADP-ribose) polymerase are
RT   differentially required for the recognition of DNA breaks and nicks
RT   and the consequent enzyme activation. Other structures recognize
RT   intact DNA.";
RL   J. Biol. Chem. 265:21907-21913(1990).
RN   [12]
RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=91035460; PubMed=2121735;
RA   Simonin F., Menissier de Murcia J., Poch O., Muller S., Gradwohl G.,
RA   Molinete M., Penning C., Keith G., de Murcia G.M.;
RT   "Expression and site-directed mutagenesis of the catalytic domain of
RT   human poly(ADP-ribose)polymerase in Escherichia coli. Lysine 893 is
RT   critical for activity.";
RL   J. Biol. Chem. 265:19249-19256(1990).
RN   [13]
RP   NUCLEAR LOCALIZATION SIGNAL.
RX   MEDLINE=92371433; PubMed=1505517;
RA   Schreiber V., Molinete M., Boeuf H., de Murcia G.M.,
RA   Menissier de Murcia J.;
RT   "The human poly(ADP-ribose) polymerase nuclear localization signal is
RT   a bipartite element functionally separate from DNA binding and
RT   catalytic activity.";
RL   EMBO J. 11:3263-3269(1992).
RN   [14]
RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=97461532; PubMed=9315851;
RA   Rolli V., O'Farrell M., Menissier-de Murcia J., de Murcia G.M.;
RT   "Random mutagenesis of the poly(ADP-ribose) polymerase catalytic
RT   domain reveals amino acids involved in polymer branching.";
RL   Biochemistry 36:12147-12154(1997).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
```

```
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-RIBOSYL)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)(N)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X16674; CAA34663.1; -.
DR   EMBL; M18112; AAA60137.1; -.
DR   EMBL; J03473; AAB59447.1; ALT_SEQ.
DR   EMBL; M17081; AAA51599.1; -.
DR   EMBL; M32721; AAA60155.1; -.
DR   EMBL; M29786; AAA51663.1; -.
DR   EMBL; M29545; AAA51663.1; JOINED.
DR   EMBL; M29766; AAA51663.1; JOINED.
DR   EMBL; M29767; AAA51663.1; JOINED.
DR   EMBL; M29768; AAA51663.1; JOINED.
DR   EMBL; M29769; AAA51663.1; JOINED.
DR   EMBL; M29770; AAA51663.1; JOINED.
DR   EMBL; M29544; AAA51663.1; JOINED.
DR   EMBL; M29782; AAA51663.1; JOINED.
DR   EMBL; X56140; CAA39606.1; -.
DR   EMBL; X56141; CAA39606.1; JOINED.
DR   EMBL; M60436; AAA60000.1; -.
DR   EMBL; M29771; AAA51663.1; JOINED.
DR   EMBL; M29772; AAA51663.1; JOINED.
DR   EMBL; M29773; AAA51663.1; JOINED.
DR   EMBL; M29774; AAA51663.1; JOINED.
DR   EMBL; M29775; AAA51663.1; JOINED.
DR   EMBL; M29776; AAA51663.1; JOINED.
DR   EMBL; M29777; AAA51663.1; JOINED.
DR   EMBL; M29778; AAA51663.1; JOINED.
DR   EMBL; M29779; AAA51663.1; JOINED.
DR   EMBL; M29780; AAA51663.1; JOINED.
DR   EMBL; M29781; AAA51663.1; JOINED.
DR   EMBL; M29783; AAA51663.1; JOINED.
DR   EMBL; M29784; AAA51663.1; JOINED.
DR   EMBL; M29785; AAA51663.1; JOINED.
DR   PIR; A26901; A26901.
DR   PIR; A28498; A28498.
DR   PIR; A29725; A29725.
DR   PIR; A35635; A35635.
DR   PIR; A33321; A33321.
DR   PIR; B33321; B33321.
DR   PIR; A39976; A39976.
DR   PIR; S14010; S14010.
DR   PIR; P26446; 1A26.
DR   HSSP; P26446; 1A26.
DR   Aarhus/Ghent-2DPAGE; 1620; NEPHGE.
DR   MIM; 173870; -.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
```

```
Query Match           25.4%;   Score 707.5;  DB 1;   Length 1013;
Best Local Similarity   34.0%;  Pred. No. 4.1e-44;
Matches 183;  Conservative 101;  Mismatches 216;   Indels 39;   Gaps 17;

Qy     1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          |||:|    ||:::       :      :          :    ||| |  ||
Db   492 VAPRGKSGA--ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 548

Qy    61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          : :| || |  :|  | :||:|  ||:    |:::|  ||||:: |:||||
Db   549 GKVFSATLGLVDIVKGTNSYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKED 608

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVK 175
          |:|: ||:||| :|||   | | |||  : || | |:|   ||:: | |: |  : ::|
Db   609 AIEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKLTVNPGTKSKLPK 666

Qy   176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQJIARGFEALEALEEAMK 235
          |        |:||| ||  |||:|: | | : :|||:|:|||||||
Db   667 P------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQAVS 720

Qy   236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
          | |    :||   | |  ::|| |||||: :| :| :|| :: |:| | : | |||: :|
Db   721 QGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLRG 779

Qy   296 PGEEEEKVDEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYRCPNLR 351
          |  |::   | | |    |  ::| :::| | ::| | |   :      |||||||||||
Db   780 GSDDSSK----DPIDVNYEKLKTDIKVVDRDSEEAEIIRKYKNTHATTHNAYDL-EVI 833

Qy   352 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKG 407
          ::| ||:|  |:  |  |::|||| |||| | | : ||  |||:||
Db   834 DIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKG 893

Qy   408 IYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIAR 467
          |||| | :|||      |   |  |       |   |: ||||
Db   894 IYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSVKGL 950

Qy   468 GQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI 526
          ||:  ||    : | |       |:||||           : :: :| |: | ||:|
Db   951 GKTTPDPSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEYIVYDIAQVNLKYLLKL 1004
```

```
RESULT 5
PPOL_BOVIN
ID   PPOL_BOVIN      STANDARD;       PRT;   1015 AA.
AC   P18493; Q9TS00;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90382673; PubMed=2119324;
RA   Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;
RT   "Cloning of a full-length cDNA encoding bovine thymus
RT   poly(ADP-ribose) synthetase: evolutionarily conserved segments and
RT   their potential functions.";
RL   Gene 90:249-254(1990).
RN   [2]
RP   SEQUENCE OF 647-714 AND 838-903 FROM N.A.
RX   MEDLINE=8151954; PubMed=2450019;
RA   Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,
RA   Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;
RT   "Depression in gene expression for poly(ADP-ribose) synthetase during
RT   the interferon-gamma-induced activation process of murine macrophage
RT   tumor cells.";
```

```
Query Match                 25.3%;       Score 704.5;    DB 1;    Length 1015;
Best Local Similarity      34.4%;       Pred. No. 6.9e-44;
Matches 187;  Conservative  95;  Mismatches 213;  Indels  49;  Gaps  20;

Qy     1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAE--ALRAAPADNRVIRVDPSCPFSRNPG  57
         : ||: ||      ||| :|| : :|     :    | :      |||    | —
Db   494 VGPKGKSGAAP--SKKSKGPVKEEGTNKSEKRMKLTLKGGAA-----VDPDSGLEHNAH 545

Qy    58 I--QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHF 112
         |      :||::|||| : : :   |||  : :|     |||| : ||||| : | :|
Db   546 VLEKGGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVIGSNKLEQM 605

Qy   113 TCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPV 170
         :| |::  | : : ||| :| :|| :  ::  : : ||  : :  ::|| |:  |||| —
Db   606 PSKEDAIEHFMKLYEEKTGNAWHSKN-FTKHPKKFYPLEID-YGQDEEAVKKLTVNPGTK 663

Qy   171 RTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAL 230
            ::    | ||    |    ::||||||||::||
Db   664 SKLPKP------VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEV 717

Qy   231 EEAMKNPTGDGQSLEELSSCYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQ 290
         |  | || :|  |  | :||:|||  |  |||| |||  | |||| :|| : || : —
Db   718 QQALSQGSDSHIL-DLSNRFTLIPFDFGMKKPPLLNNANSVQAKVEMLDNLLDIEVAY 776

Qy   291 TLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYR 346
         ::  :  : |            |:|    :|| : | | :| ||| :|:|     : | ||
Db   777 SLLRGGSDDSSK-----DPIDVNYEKLKTDIKVVDKDSEEAEIIRKYVKNTHATTHNAYD 831

Qy   347 CPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGG 402
            :: : ||||                || : ||||||:::      :|||| : —     :|
Db   832 L-EVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGY 890

Qy   403 RVGKGIYFASENSKSAGYVTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFD 462
         : :| ||| |   | | :|  :: || |||| | | :||:| :|  ||| | |  —
Db   891 MFGKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEAALGNMYELK-HARHISKLPKGKH 947

Qy   463 SVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRY 522
         |:    |   |  |:|  || ||   ||  |  | |  | ::| ||:|  | :| : ||
Db   948 SVKGLGKTTPDPSASI--TVDG--VEVPLGTGIS-SGVNDTCLLYNEYIVYDIAQVHLKY 1002

Qy   523 LLEI 526
         || ::
Db  1003 LLKL 1006
```

```
RESULT  6
PPO2_HUMAN
ID   PPO2_HUMAN      STANDARD;      PRT;     583 AA.
AC   Q9UGN5; Q9Y6C8; Q9NUV2; Q9UMR4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE   (hPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoeger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE OF 2-583 FROM N.A. (ISOFORM 1).
```

```
RL   Eur. J. Biochem. 171:571-575(1988).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY[ADP-RIBOSYL]ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D90073; BAA14114.1; -.
DR   EMBL; X06986; CAA30046.1; -.
DR   EMBL; X06987; CAA30047.1; -.
DR   PIR; JS0428; JS0428.
DR   PIR; S00328; S00328.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   DNA_BIND      1    374
FT   DOMAIN      375    525       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      386    462       BRCT.
FT   DOMAIN      526   1015       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     127    164       PARP-TYPE.
FT   DOMAIN      209    211       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     408    408       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     414    414       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     446    446       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     449    449       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     457    457       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     472    472       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     489    489       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     492    492       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     514    514       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     515    515       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     521    521       ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE   1015 AA;  113355 MW;  0A5FE9D9F04F5B04 CRC64;
```

```
Matches  185;  Conservative  97;  Mismatches  207;  Indels  58;  Gaps  18;

Qy    13  GSKKQRQGTEEED-------SFRSTAEALRA------APADNRVIRVDPSCPFSRNPG-  57
              |  |  ||             :  :||:   :        ||      |||
Db    54  GGKANKDRTEDKQDGMPGRSWASKRVSESVKALLLKGKAP------VDPECTAKVGKAH 106

Qy    58  --IQVHEDYDCTLNQTNIGNNNNKFYIIQLLEEGSR--PFCWNRWGRVGEVGQSKMNHFT 113
              :  :::  ||  ||||:  ||||:||||:||:::  :  |  |||||||  ::  ||
Db   107  VYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSL--VA 164

Qy   114  C---LEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPV 170
              |          |||||||:||||||       ||  ||: :||:  ::| :|:|
Db   165  CSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEETKKEESL 224

Qy   171  RTVVKPCS-LDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA 229
              ::  |||:  |||||||||:|:|||||:||||||||||||||||||  || ||::|:::
Db   225  KSPFKPESQLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKK 284

Qy   230  LEEAMKNPTGD-GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIEL 288
              :|::: |:  ||::|:  |||||||||:||| ||| :| |||||| ||||::|||
Db   285  IEDCIR--AGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEI 342

Qy   289  AQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSY 345
              ::||::   ||| ||||||||||||:||:||:| |||| |:||||::|||::      ||
Db   343  AIKL-----VKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSD 396

Qy   346  RCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SG 401
              :|  |::     ||::|   |:|||||||||||||  |:||:||||||||||:     :|
Db   397  YTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITG 454

Qy   402  GRVGKGIYFASENSKSAGIYVTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGF 461
              |||||||      |:  ||||    ||   :  |||||::||  |||  ||:
Db   455  YMFGKGIYFADMSSKSANYCFASRL--KNTGLLLLSEVALGQCNELLEANPKAEGLLQGK 512

Qy   462  DSVIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSF--SQSEYLIYKESQCR 519
              |::||| :  |||| ||  :|||  ||||||||||||||||||      :|::|||:|||
Db   513  HSTKGLGKMAPSSAHFV--TLNGS--TVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVR 568

Qy   520  LRYLLEI 526
              :|:|||
Db   569  MRYLLKV 575
```

```
RESULT
PPOL_RAT
7
ID   PPOL_RAT              STANDARD;        PRT;     1013 AA.
AC   P27008; O35937;
DT   01-AUG-1992  (Rel.  23,  Created)
DT   15-JUL-1998  (Rel.  36,  Last sequence update)
DT   16-OCT-2001  (Rel.  40,  Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Monocytes;
RX   MEDLINE=98046546; PubMed=9385436;
RA   Beneke S., Meyer R., Buerkle A.;
RT   "Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly
RT   (ADP-ribose) polymerase.";
RL   Biochem. Mol. Biol. Int. 43:755-761(1997).
RN   [2]
RP   REVISION TO 811.
RA   Beneke S., Meyer R., Buerkle A.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-11 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Prostate;
RX   MEDLINE=92290013; PubMed=1601134;
```

RC   TISSUE=Fetal brain;
RX   MEDLINE=99263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [3]
RP   SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fibroblast;
RX   MEDLINE=99268466; PubMed=10038144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian poly(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA   Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii Y., Kawai Y.,
RA   Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA   Masuho Y., Kanehori K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC   SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (By
CC   similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC   produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN
CC   THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO
CC   DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;
CC   LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND
CC   THYMUS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AJ236912; CAB65088.1; -.
DR   EMBL; AF085734; AAD29857.1; ALT_INIT.
DR   EMBL; AJ236876; CAB41505.2; ALT_INIT.
DR   EMBL; AK001980; BAA92017.1; ALT_TERM.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Alternative splicing.
FT   DNA_BIND   1   88   POTENTIAL.
FT   DOMAIN   86   583   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   4   7   NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT   (POTENTIAL).
FT   DOMAIN   35   40   NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT   (POTENTIAL).
FT   VARSPLIC   68   80   MISSING (IN ISOFORM 2).
FT   CONFLICT   447   447   P -> H (IN REF. 2).
FT   CONFLICT   481   481   N -> H (IN REF. 4).
SQ   SEQUENCE   583 AA; 66205 MW; 5B7AE8AE531836AF CRC64;

Query Match   25.1%; Score 699; DB 1; Length 583;
Best Local Similarity 33.8%; Pred. No. 8.1e-44;

```
RA   Potvin F., Thibodeau J., Kirkland J.B., Dandenault B.,
RA   Duchaine C., Poirier G.G.;
RT   "Structural analysis of the putative regulatory region of the rat
RT   gene encoding poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 302:269-273(1992).
RN   [4]
RP   SEQUENCE OF 514-1013 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=90027702; PubMed=250873l;
RA   Thibodeau J., Gradwohl G., Dumas C., Clairoux-Moreau S., Brunet G.;
RT   "Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase
RT   catalytic domain and analysis of mRNA levels during the cell cycle.";
RL   Biochem. Cell Biol. 67:653-660(1989).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND RECOVERY OF CELL FROM DNA DAMAGE.
CC       EVENTS INVOLVED IN THE REGULATION OF THE MOLECULAR
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U94340; AAC53544.1; -.
DR   EMBL; X65496; CAA46477.1; -.
DR   EMBL; X65497; CAA46478.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
DR   PROSITE; PS00347; BRCT; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0
FT   DNA_BIND      1    372
FT   DOMAIN      385    461    BRCT.
FT   DOMAIN      373    523    AUTOMODIFICATION DOMAIN.
FT   DOMAIN      524   1013    NAD-BINDING.
FT   ZN_FING      20     55    PARP-TYPE.
FT   ZN_FING     124    161    PARP-TYPE.
FT   DOMAIN      206    208    NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225    NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES       1      1    BLOCKED (BY SIMILARITY).
FT   MOD_RES     407    407    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     413    413    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     435    435    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     437    437    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     456    456    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     484    484    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     488    488    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     491    491    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512    ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     519    519    ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    638    638    Y -> H (IN REF. 4).
FT   CONFLICT    641    641    E -> A (IN REF. 4).
FT   CONFLICT    752    752    N -> D (IN REF. 4).
SQ   SEQUENCE   1013 AA;  112529 MW;  AA566F2B29BE97C0 CRC64;

Query Match             25.0%;  Score 698;  DB 1;  Length 1013;
Best Local Similarity   33.6%;  Pred. No. 2.1e-43;
  Matches 182;  Conservative 101;  Mismatches 214;  Indels 44;  Gaps 18;

Qy    1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
         : || ||  : :|:  :  :|:       :|:      |:    :   ::  | | :|
Db  493 VVPKGKSAAPSKKKSK----GAVKEEGVNKSEKRMKLITLKGGAAVDPDSGLEHSAHVLEKG 548

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLE--EGSRFFCWNRWRGRVGEV-GQSKMNHFTCLED 117
         | ||:   |||   |  ||:|:|  ||  | ||| : |:| :|| |  |:|
Db  549 GKVFSATGLVDLVKGTNSYYKLQLLESDKESRYWIFRSWGRVGTVIGSNKLEQMPSKED 608

Qy  118 AKKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKP- 176
         ||| ||| || |||:  |   |:  ||  :|::|  :  :  :|
Db  609 AVEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL------AVKPG 659

Qy  177 --CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234
           : ||::|:|    :|:| :  ::|:  | :  :||||||| |:|  : : : :|:
Db  660 TKSKLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQAV 719

Qy  235 KNPTGDCGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
         :  :  :|  |  :::| |:|  |  |  | |  |: ||  |:  :: :|: ||  |
Db  720 SQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNTDSVQAKVEMLDNLLDIEVAYSLLR 778

Qy  295 APGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNL 350
         :          : | :  |:|: :||:: :  :|::| :|: |:    ||| | |:|
Db  779 GGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDL-EV 832

Qy  351 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK 406
         : : |||: ||  : :||||:| :::|||||:   ||   | |  |:    ||| |||
Db  833 IDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTNFAGILSQGLRIAPPEAPVTGYMFGK 892

Qy  407 GIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIA 466
         |||||:|| : :   |::   : || | |:|| ||| :  ||  |:|  ||  |  ||
Db  893 GIYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSVKG 949

Qy  467 RGQTEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRYLLE 525
         |: : | |  ::||| | |:    | | |     | ||:| :|| :| ::| | ||
Db  950 LGKTAPDPSASI--TLDG--VEVPLGTGI--PSGVNDTCLLNEYIVYDIAQVNLKYLLK 1003

Qy  526 I 526

Db 1004 L 1004

RESULT 8
PPOL_ARATH
ID   PPOL_ARATH      STANDARD;      PRT;    637 AA.
AC   Q11207;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   APP.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
Db   459 QLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKG 518
             ::: .|  | |||| |       ||  |||||||||  :   ||    :|  |||
Qy   408 IYFASENSKSAGYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSV 464
             :||| |||| ||       | : ||||| |     ||||| :    :   |||  |
Db   519 VYFADMFSKSANY-----CYANTGANDGVLLCEVALGDMNELLYSDYNADNLPPGKLST 573
             ||:|| :|||| |||    ||| ||  ||  :|  |: :|||||||| |::|:
Qy   465 IARGQTEPDAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLL 524
             |:|: ||  |  |  || :|: |||  |||  ||::|:
Db   574 KGVGKTAPNPSEAQTLE-DG--VVPLGKPVE-RSCSKGMLLYNEYIVYNVEQIKMRYVI 629

Qy   525 EI 526
             ::
Db   630 QV 631
```

RESULT   9
PPOL_CHICK

```
ID   PPOL_CHICK           STANDARD;          PRT;    1011 AA.
AC   P26446;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=91340148; PubMed=1840535;
RA   Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;
RT   "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid
RT   sequence and comparison with mammalian enzyme sequences.";
RL   Gene 102:157-164(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.
RX   MEDLINE=96353841; PubMed=8755499;
RA   Ruf A., Mennissier de Murcia J., de Murcia G.M., Schulz G.E.;
RT   "Structure of the catalytic fragment of poly(AD-ribose) polymerase
RT   from chicken.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO
RP   895.
RX   MEDLINE=98191351; PubMed=9521710;
RA   Ruf A., de Murcia G.M., de Murcia G.M., Schulz G.E.;
RT   "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived
RT   from crystal structures and homology modeling.";
RL   Biochemistry 37:3893-3900(1998).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.
RX   MEDLINE=98239716; PubMed=9571033;
RA   Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;
RT   "The mechanism of the elongation and branching reaction of poly(ADP-
RT   ribose) polymerase as derived from crystal structures and
RT   mutagenesis.";
RL   J. Mol. Biol. 278:57-65(1998).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
```

```
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95269779; PubMed=7750552;
RA   Lepiniec L., Babiychuk E., Kushnir S., van Montagu M., Inze D.;
RT   "Characterization of an Arabidopsis thaliana cDNA homologue to animal
RT   poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 364:103-108(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z48243; CAA88288.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND      1    140       POTENTIAL.
FT   DOMAIN      140    637       NAD-BINDING (BY SIMILARITY).
FT   DOMAIN       41     62       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   637 AA;  72175 MW;  527A8F46460SD127 CRC64;
```

```
Query Match             25.0%;  Score 697.5;  DB 1;  Length 637;
Best Local Similarity   34.9%;  Pred. No. 1.2e-43;
Matches  189;  Conservative  91;  Mismatches 219;  Indels  43;  Gaps  17;

Qy     6 KASVQTEGSKKQRQGTEEEDSFRS--TAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED 63
           |:| :|:  :|  |  |||| |     |   | |::  |  ||::|:   |  || :|
Db   112 KSSNGTDEAEDDNNGFEEEKKEEKIVTATKKGAAVLDQWI----PDEIKSQYHVLQRGDD 167
            |: : ||||||||::|||||:::|| :||    || |:   ||||  ::::||:   ||
Qy    64 -YDCTLNQTNIGNNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMN-HFTCLEDAK 119
            | |:||||:||||||| ::::||::||    | |||||:|| :|  ||  |: | |||
Db   168 VYDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVYTRWGRVGVKGQSKLDGPYDSWDRAI 227
           | : :||:||| ||||| :| |:||  :  ::    ||| ||: |  |:||    |  |:
Qy   120 KDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSSPVRTVVKP--C 177
            | | | || :| ::  :|:|  :|   |  |     | : :  :|  |  |: |     |
Db   228 EIFTNKFNDKTKNKWNWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNNDIPSSSSEVKPEQS 287
            |:  :| :|:|||| : :|| |:|    || | |      |  | |   |  | |||   |
Qy   178 SLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNP 237
           |  |  |||:::|  | ||    |  |  | |||||:|  :|  :|   | |  | :|| |
Db   288 KLDTRVAKFISLICNVSMAQHMEIGYNANKLPLGKISKSTISKGYEVLKRISEVIDR- 346
           |||:|| ||:|:|  | : |  || :|:|||||| |  ||:|||||:|:|:||:|||
Qy   238 TGDGQSLEELSSCFYTVIPHNFGRSRPP--INSPDVLQAKKDMLLVLADIELA-QTLQA 294
            ||| | ||| |:: :| |: |    |   :| |  |:||  ::|:  |  ||  ||| |
Db   347 -YDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLLSV 405
           |  :| ||| |  |::|| | |      | ||| || ||:|:  | | :|:||| |||::|
Qy   295 APGEEEKVEEVPHLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRCPNLR 351
            |  |||||  |  | | :||| |  || :  | |  |  || |:      ||||:| |
Db   406 DPGLQDD------PLYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVEIA 458
           ||| | |      ||:| |::  | | | |:| ::::  | | ::  :| |:|: |:|
Qy   352 HVKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKG 407
```

```
Db         :  |:::: |  ::|| ::|::: |              |
        497 KPANMKSAGKVKEEQGPSKSEKKMKLTVKGGAAVDPDSGL--EDSAHVFEKGGKI-----  549

Qy         :  |::: | :|  :|||::::|:|::::|::: |:|           :|:|
         64 YDCTLNQTIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKK   120

Db         :  || : |:|:||||:|:: ||:|: : :|::                 |||
        550 FSATLGLVDIVKGTNSYYKLQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAVE  609

Qy     121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVKPCS  178

Db         :  |::|: |:|:  ||::|:: |  :|:: |:||               :||
        610 HFLNLYEEKTGNSWHSKN-FTKYPKKFYPLEID-YGQDEAVRKLTVSAGTKSKLAKP--   665

Qy     179 LDPATQNLJTNIFSKEMFKNAMTLMNLDVKKMPLGKLIKQQIARGFEALEALEEAMKNPT  238

Db         :  :|:| |||||:||||:||||||
        666 ----IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSLLNEVQQAVSDGG  721

Qy     239 GDGQSLEELSSCFVTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGE  298

Db         :  |: |::||:||:|| :: :: || : |||:| :::|: || :|
        722 SESQIL-DLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRGGNE  780

Qy     299 EEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLRHVW   354

Db         :  : :|  | |:  |:: |:|||:| ||| :|: ||
        781 DGDK-----DPIDINYEKLRTDIKVVDKDSEEAKIIKQYVKNTHAATHNAYDL-KVVEIF  834

Qy     355 KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYF  410

Db         :  :: |:|| :|||| || :::|||||  : |: |||| :||:||
        835 RIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYF  894

Qy     411 ASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEH------HITIDDPSLKSPPPGFDSV  464

Db         :  | |||||| : |:|:|:|: ||||
        895 ADMVSKSANYCHTSQ--ADPIGILLLGEVALGNMYELKNASHIT-------KLPKGKHSV  945

Qy     465 IARGQTEPDPAQDIELELDGQPVVVPQGPVQCPSFKSSSFSQSEYLIYKESQCRLRYLL   524

Db         :  || : || |:|  |::|:|| : :|||||| ::
        946 KGLGKTAPDPT--ATTTLDG--VEVPLGNGIS-TGINDfCLLYNEYIVDVAQVNLKYLL  1000

Qy     525 EI 526

Db         ::
       1001 KL 1002


RESULT 10
PPOL_MOUSE
ID   PPOL_MOUSE     STANDARD;      PRT;   1012 AA.
AC   P11103; Q9JLX4; Q9VQ3;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1) (mSPARP).
GN   ADPRT OR ADPRT1 OR ADPRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=BXSB;
RX   MEDLINE=89263780; PubMed=2498841;
RA   Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;
RT   "Sequence and organization of the mouse poly (ADP-ribose) polymerase
RT   gene.";
RL   Nucleic Acids Res. 17:3387-3401(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC   STRAIN=129/Sv C57BL/6; TISSUE=Fibroblast;
RX   MEDLINE=20270268; PubMed=10809783;
RA   Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;
RT   "Characterization of sPARP-1. An alternative product of PARP-1 gene
RT   with poly(ADP-ribose) polymerase activity independent of DNA strand
RT   breaks.";
RL   J. Biol. Chem. 275:15504-15511(2000).
RN   [3]
RP   KNOCK-OUT.
```

```
CC   -!-  MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC        AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC        FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC        THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC        AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!-  SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!-  SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```
```
DR   EMBL; X52690; CAA36917.1; -.
DR   PIR; JH0581; JH0581.
DR   PDB; 2PAW; 27-MAY-98.
DR   PDB; 2PAX; 15-MAY-97.
DR   PDB; 1PAX; 27-MAY-98.
DR   PDB; 2PAX; 27-MAY-98.
DR   PDB; 3PAX; 27-MAY-98.
DR   PDB; 4PAX; 27-MAY-98.
DR   PDB; 1A26; 27-MAY-98.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; zinc-finger; zinc; 3D-structure.
FT   DNA_BIND        1    370
FT   DOMAIN        371    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN        382    458       BRCT.
FT   DOMAIN        523   1011       NAD-BINDING.
FT   ZN_FING        21     56       PARP-TYPE.
FT   ZN_FING       125    162       PARP-TYPE.
FT   DOMAIN        207    209       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN        220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES       403    403       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       404    404       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       410    410       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       411    411       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       432    432       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       441    441       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       442    442       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       453    453       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       454    454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       468    468       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       481    481       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       485    485       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       509    509       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       510    510       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       517    517       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT      895    895       A -> R (IN REF. 1).
SQ   SEQUENCE   1011 AA;  113520 MW;  261AED9383139144 CRC64;

Query Match      24.9%;  Score 694;  DB 1;  Length 1011;
Best Local Similarity  34.1%;  Pred. No. 4e-43;
Matches 185;  Conservative 97;  Mismatches 204;  Indels 56;  Gaps 20;

Qy    ,    5 RKASVQTEGSKKQRQG-TEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED  63
```

```
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jaeger S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC       short form/sPARP-1: may be produced by alternative initiation.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```

```
DR   EMBL; X14206; CAA32421.1; -.
DR   EMBL; AF126717; AAF61293.1; ALT_INIT.
DR   PIR; S04200; S04200.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1340806; Adprt1.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Zinc; Alternative initiation.
KW   ADP-ribosylation; Zinc-finger; Nuclear protein;
FT   INIT_MET      0      0
FT   CHAIN         1   1012       POLY [ADP-RIBOSE] POLYMERASE-1, LONG
FT                                ISOFORM.
FT   CHAIN       521   1012       POLY [ADP-RIBOSE] POLYMERASE-1, SHORT
FT                                ISOFORM.
FT   INIT_MET    521    521       FOR SHORT ISOFORM.
FT   DNA_BIND      1    371
FT   DOMAIN      372    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      384    460       BRCT.
FT   DOMAIN      523   1012       PARP-TYPE.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       BY SIMILARITY.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     406    406       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     412    412       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     443    443       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     447    447       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     455    455       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     483    483       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     487    487       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     490    490       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     511    511       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N]    (POTENTIAL).
FT   MOD_RES     518    518       ADP-RIBOSYL[N]    (POTENTIAL).
FT   CONFLICT    590    590       L -> V (IN REF. 2).
FT   CONFLICT    607    607       E -> D (IN REF. 2).
FT   CONFLICT    611    611       Q -> H (IN REF. 2).
FT   CONFLICT    628    628       N -> D (IN REF. 2).
FT   CONFLICT    678    678       D -> E (IN REF. 3).
FT   CONFLICT    702    702       R -> F (IN REF. 3).
FT   CONFLICT    716    716       Q -> E (IN REF. 2).
FT   CONFLICT    757    757       Q -> L (IN REF. 3).
FT   CONFLICT    856    856       R -> F (IN REF. 3).
FT   CONFLICT    981    981       A -> C (IN REF. 2).
SQ   SEQUENCE  1012 AA;  112968 MW;  4354C3E5F01B9439 CRC64;
```

```
Query Match
Best Local Similarity  24.7%;              Score 688;  DB 1;  Length 1012;
                       32.7%;              Pred. No. 1.1e-42;
Matches 177;  Conservative 106;  Mismatches 214;  Indels 44;  Gaps 17;

Qy     1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
           ||:|.|:::.:::|                   :|                :|
Db   492 VAPRGKSAAPSKKSK----GCFKEEGVNKSEKRMKITLKGGAAVDPDSGLEHSAHVLEKG 547

Qy    61 HEDYDCTLNQTNIGNNNKFYIIQLLEEG--SRFCWNRWGRVGEV-GQSKMNHFTCLED 117
           || :||||  : :|||||:|||||| |||:||  :||||||||
Db   548 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRLGTVIGSNKLEQMPSKEE 607

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKP- 176
           |:|.:|| :|| :::|  :|:  ::|||: |:|| : |||  :::::|||
Db   608 AVEQFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL------TVKPG 658

Qy   177 --CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234
             .||.| :|:|||||| : :||:.||:|:|:|:||||||| :|||   :|:::|::
Db   659 TKSKLPKPVQELVGMIFDVDSMKKALVEYEIDLQMPLGKLSRRQIQAAYSILSEVQQPV 718

Qy   235 KNPTGDGQSLEELLSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
           ||  |  | ||| ||:||||:||| ||::|||:::.:|||| :|:| :|||||:|| ||
Db   719 SQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLR 777

Qy   295 APGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYRCPNL 350
            |:::: : .:|:|:: :: :::: :::||::|||:|| ::|:|:    ||:|:||:|:
Db   778 GGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDL-EV 831

Qy   351 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK 406
            :::::||||| :|::: ||:||||||||| : |: :|:|:|||| .|    :|| :|||
Db   832 IDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGK 891

Qy   407 GIYFASENSKSAGYVTTMHCGGHGVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIA 466
           |||||.:.||||.| :: :  |: :|:::|||||||::::||  :|| |||| :|| :|
Db   892 GIYFADMVSKSANYCHTSQ--GDPIGLIMLGEVALGNMYELK-HASHISKLPKGKHSVKG 948

Qy   467 RGQTEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRYLLE 525
            |:||||::: | ||:   |:|:||:|: : |  | :::|: ::|::| | :::|:|||:
Db   949 LGKKTPDPSASITLE----GVEVPLGTGI-PSGVNDTALLYNEYIVYDIAQVNLKYLLK 1002

Qy   526 I 526
           |
Db  1003 L 1003

RESULT 11
PPOL_XENLA
ID   PPOL_XENLA      STANDARD;      PRT;   998 AA.
AC   P31669;
```

```
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Saulier-Le Drean B.M.;
RL   Thesis (1992), University of Rennes, France.
RN   [2]
RP   SEQUENCE OF 742-876 FROM N.A.
RX   MEDLINE=93277538; PubMed=8503897;
RA   Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,
RA   Miwa M.;
RT   "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)
RT   polymerase from Xenopus laevis and cherry salmon using heterologous
RT   oligonucleotide consensus sequences.";
RL   Biochem. Biophys. Res. Commun. 193:119-125(1993).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,
CC       AND BRAIN. LOW IN LIVER.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; Z12139; CAA78126.1; -.
DR   EMBL; D13810; BAA02966.1; -.
DR   PIR; S31735; S31735.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   NON_TER       1      1
FT   DNA_BIND     <1    356
FT   DOMAIN      357    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      369    445       BRCT.
FT   DOMAIN      508    998       NAD-BINDING.
FT   ZN_FING       8     43       PARP-TYPE.
FT   ZN_FING     111    148       PARP-TYPE.
FT   DOMAIN      193    195       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      207    212       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     391    391       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     397    397       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     419    419       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     428    428       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     429    429       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     447    447       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     454    454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     467    467       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     471    471       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     477    477       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     495    495       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     496    496       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     503    503       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    746    746       Q -> E (IN REF. 2).
SQ   SEQUENCE   998 AA;  111126 MW;  F5A25E4A3366BAE7 CRC64;

Query Match             24.4%;  Score 681;  DB 1;  Length 998;
Best Local Similarity   33.6%;  Pred. No. 3.6e-42;
Matches  178;  Conservative   95;  Mismatches  215;  Indels   42;  Gaps   18;

Qy         10  QTEGSKKQRQGT-EEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHEDYDCTL   68
                  ::   :::   : :: : :: ::::        : ||     ||   :  : ||
Db        488  KSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDS---ELEDSCHVLETGG----KIFSATL  540

Qy         69  NQTNIGNNNNKFYIIQLLE--EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKKDFKKK  125
                :: : :: ::  : : |||  ::  : ||:  : ||: || : :  ::  || :  : :
Db        541  GLVDITRGTNSYYKLQLIEHDRSRYWFRSWGRVGTVIGSKKLEEMSSKEDAIEHFLNL   600

Qy        126  FWEKTKNKWEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKVDSGPVRTVVKPCSLDPATQ  184
                |: : ::  : :: ::   :|:  :   ::   ::::::  :::  :: : ::   :
Db        601  YQDKTGNAWHS-PNFTKYPKKFYPLEIDYGQEEDVVKKLSVGAGTKSKLAKP-----VQ  653

Qy        185  NLITNIFSKEMPKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEAMKNPTGDGQSL  244
                ::|   :  : :  : :::: :: ||| ::: |::   :   ::: : :   :  :
Db        654  ELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAVSESLSEARLL  713

Qy        245  EELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTIQAAPGEEEKVE   304
                : ||  :|: ||:||:   | : :: ||| || | : :  : || : || ||
Db        714  -DLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDGEK--  770

Qy        305  EVPHPLDRDYQLLRCQLQLLDSGESEYKAIQYLK----QTGNSYRCPNLRHVWKVNREG  360
                :||| :  :   :::   ::: :::  : ::        :|||: | || :  :  ::|
Db        771  ---DPIDVKYEKIKTDIKVAKDSEESRIICDYVKNTHADTHNAYDLEVL-EIFKIDREG  826

Qy        361  EGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSK  416
                ||||  : :|  ||| ||:  :  :  :: : :||: || : ||| :| || :|||||
Db        827  EYQRYKPFKQLNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSK  886

Qy        417  SAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQ  476
                |||||| : |   | |:|| ||:| |||::|| |:  |: : : | ||  : |  ||||
Db        887  SANYCHAM--PGSPIGLILLGEVALGNMHELKAASQITKL-PKGKHSVKGLGRTAPDPSA  943

Qy        477  DIELELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEI  526
                ::||:  |   :| ::      |::  |:: :::  :  :: | | ::|
Db        944  TV--QLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLKYLLKL  988


RESULT 12
PPOL_SARPE
ID   PPOL_SARPE     STANDARD;      PRT;   996 AA.
AC   Q11208;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
```

```
Qy    22  EEEDSFRSTAEALRAAPADNRV------IRVDPSCPF---------SRNPGIQVHEDYDCTL  68
          || ||  :  :: ||:: |||                          |||  ||:  |
Db   489  EESKSKSKSIYTKSVPKSMTLKIKDGLAVDDPDSGLEDVAHVYVSRN----KEKYNVVL 543

Qy    69  NQTNIGNNNKFYIIQLLEEG--SRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKKK 125
          ||   ::  |  : ||:|||     || ||  || ||| |:: |::| :  ||  ||:
Db   544  GITDIQKNKNSFYKLQLLESDMKNRFWFRSWGRIGTTIGGNKIDNFSNLVDAIVQFKEL 603

Qy   126  FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLDPATQN 185
          :  |||  :  ||| |  :   |     | ::         ||    |  : |
Db   604  YLEKSGNHFENRENFVKVAGRMYPIDIDYAEDS-----KIDLSAEHDIKSKLPL--SVQD 656

Qy   186  LITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQOIARGFEALEALEEAMKNPTGDGQSLE 245
          :  | :| ||    :||||||:||||   |||:|  | ||: |  |
Db   657  IIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKIQSAYKVLTEIYELIQG-GGTNAKFI 715

Qy   246  ELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVEE 305
          :: ||  |   |     || :  |  ||   | :|   |  |  ||::::    |
Db   716  DATNRFYTLIPHNFGTQSPPLLDTTEQVEQLRQMLDSLIEIECAYSLLQT---EDSKAD- 771

Qy   306  VPHPLDRDYQLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKVNREGEG 362
          :: |: |: || |  ||   |   ||      | |    |||      |||||| |||
Db   772  -INPIDKHYEQLKTKLEPLDKNSEEYILLQKYVKNTHAETHKLYDLEVVDIFKVARQGEA 830

Qy   363  DRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH-----SGGRVGKGIYFASENSKSA 418
          || |    |||||||  |  :  | ||:| |||   |:| ||| ||||||||  ||
Db   831  RRYKPFFKLHNRRLLWHGSRLTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADMVSKSA 890

Qy   419  GVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDI 478
          || |  |   |||||| |  |     | |||  : |  :||         ||  | |
Db   891  NYCCTSH--HNSTGLMLLSEVALGDMMECTAAKVTKLPNDK-HSCFGRGRTMPNPSESI 947

Qy   479  ELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIH 527
          :| |  | || || |||| | ::|||    ||  |  :|:   : :| ::
Db   948  IRE-DG--VEIPLGKPITNDSLK-SSLLYNEFIYDIAQVNIQYMLRMN 992
```

RESULT 13
PPOL_DROME
ID   PPOL_DROME   STANDARD;   PRT;   994 AA.
AC   P35875; Q9W5Q5; Q9W5S1;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   PARP OR CG17696/CG17718.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93234521; PubMed=8475096;
RA   Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,
RA   Sugimura T., Miwa M.;
RT   "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:
RT   leucine zipper in the auto-modification domain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND
RP   TISSUE SPECIFICITY.
RC   STRAIN=CANTON-S;
RX   MEDLINE=98234380; PubMed=9555614;
RA   Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;
RT   "Genomic organization of Drosophila poly(ADP-ribose) polymerase and
RT   distribution of its mRNA during development.";
RL   J. Biol. Chem. 273:11881-11886(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (LONG ISOFORM)
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;

DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
OS   Sarcophaga peregrina (Flesh fly) (Boettcheriscia peregrina).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7386;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=94170813; PubMed=8125121;
RA   Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,
RA   de Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;
RT   "Cloning and functional expression of poly(ADP-ribose) polymerase
RT   cDNA from Sarcophaga peregrina.";
RL   Eur. J. Biochem. 220:607-614(1994).
-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; D16482; BAA03943.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; FALSE_NEG.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   DNA_BIND   1   369   BY SIMILARITY.
FT   DOMAIN   370   507   AUTOMODIFICATION DOMAIN.
FT   DOMAIN   382   456   BRCT.
FT   DOMAIN   508   996   NAD-BINDING.
FT   ZN_FING   19   54   BY SIMILARITY.
FT   ZN_FING   126   164   BY SIMILARITY.
FT   DOMAIN   211   214   NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN   232   235   NUCLEAR LOCALIZATION SIGNAL 2ND PART.
SQ   SEQUENCE   996 AA; 113018 MW; 69DDD36F7487298 CRC64;

Query Match   23.4%; Score 652; DB 1; Length 996;
Best Local Similarity   33.3%;
Matches 176; Conservative 100; Mismatches 205; Indels 48; Gaps 18;

Pred. No. 4.8e-40;

```
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES
CC       OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM
CC       IN LATER EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN
CC       EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```

```
DR   EMBL; D13806; BAA02964.1; -.
DR   EMBL; AF051548; AAC24518.1; -.
DR   EMBL; AF051544; AAC24518.1; JOINED.
DR   EMBL; AF051545; AAC24518.1; JOINED.
DR   EMBL; AF051546; AAC24518.1; JOINED.
DR   EMBL; AF051547; AAC24518.1; JOINED.
DR   EMBL; AE002935; AAF45400.1; -.
DR   EMBL; AE002666; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE002892; AAF45445.2; ALT_SEQ.
DR   PIR; A47474; A47474.
DR   HSSP; P26446; 1A26.
DR   FlyBase; FBgn0010247; Parp.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP_reg.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; PARP_ZN_FINGER_1; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.
FT   DNA_BIND       1    367
FT   DOMAIN       368    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN       380    454       BRCT.
FT   DOMAIN       508    994       NAD-BINDING.
FT   ZN_FING       19     54       PARP-TYPE.
FT   ZN_FING      123    161       PARP-TYPE.
FT   DOMAIN       208    210       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN       223    228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   VARSPLIC     376    564       MISSING (IN SHORT ISOFORM)
SQ   SEQUENCE   994 AA;   113791 MW;  ACA85A270DD29E08 CRC64;

Query Match          22.3%;
Best Local Similarity  31.6%;    Score 621.5;  DB 1;  Length 994;
Matches 167;   Conservative 102;  Mismatches 221;   Indels 39;   Gaps 17;

Qy   16 KQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQ---VHED----YDCT  67
          : :||              :  ::    |  :       :  ::     |: :
Db  481 KSRIPKET=KSLNSNSIYTKSMPV-SRTFKVKDGLAVDPDSGLEDIAHVVDSNNKYSVV 539

Qy   68 LNQINIGNNNNKFYIIQLL--EEGSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKK 124
          |:|: | | :|||    ||| : :  ||:||:: |::||||  :::||::
Db  540 LGLTDIQRNKNSYYKVQLLKADKPKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKE 599

Qy  125 KFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLDPATQ 184
          |  : :||||    :|  ::| |  ||:   ||:  :  |  |  |:
Db  600 IYADKTGNEYEQRDNFVKRTGRMYPIEIQ--YDDDQKLVKHESHFFTS----KLEISVQ 652

Qy  185 NLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEEAMKNPTGDGQSL 244
          ||| ||     ||: :  | ||||||||||:|| :||||||  |||||
Db  653 NLIKLIFDIDSMNKTLMEFHIDMDKMPLGKLSAHQIQSAYRVVKEIYNVLECGSNTAK-L 711

Qy  245 EELSSCFYIVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEEKVE 304
          : :  ||  | : || : : ||||  |: :|:|| :| | : | :::||
Db  712 IDATNRFYTLIPHNFGVQLPTLIETHQQIEDLRQMLDSLAEIEVAYSI----IKSEDVS 766

Qy  305 EVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKVNREGE 361
          : :||  : ||:| : | |  :::  |:| ||  ||| |||::| ||:  | |:|:||
Db  767 DACNPLDNHYAQIKTQLVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKVSRQGE 826

Qy  362 GDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRI--MPHSGGRVGKGIYFASENSKS 417
          || | ||  |  :|||||| | : ::|| || :  | ||||||||||||
Db  827 ARRFKPFKKLHNRKLLWHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFADMVSKS 886

Qy  418 AGYVTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQD 477
          ||  ||: |:| ||:||||||:| || |:|:: |   ||   :| | |  |:|  |
Db  887 ANYCCTSQ--QNSTGLMLLSEVALGDMMECT-SAKYINKLSNNKHSCFGRGRTMPDPTKS 943
```

```
Qy  358 ------------------------REGEGDRFQAHSKLG--------NRRLLWHGTNVAVVAAILTS 392
                                  :: ||  |  | :|| :|  ::|||
Db  340 FELKDPNEIPTEAPVEQEVPKKRGRKSTKTAAPTVPPPTKRLLWHGTRVTNVFSILMN 399

Qy  393 GLR--IMPHSGGRVGKGIYFASENSKSAGVTTMHC--GGHQVGYMFLGEVALGK----- 443
            ||      |   |  :||||   : :| | ||| :|     ||:  :| ||
Db  400 GLQFPVGDRCGLMFGNGVYFANVPTKSANYC----CPEASKRV-FMLLCEVETANPLVLY 454

Qy  444 EHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKS- 502
            | |:              :|    :|                       ||:|
Db  455 ESEIDADEKMEKAKK---TSVYAAGKHTPRDT-----VEING--------IPAFKSN 495

Qy  503 -------SSFSQSEYLIYKESQCRLRYLLEI 526
               :    |:||:    :|:|:||
Db  496 LETIEEETRLLYDEYVMFNKEHFKIKYVVEV 526
```

RESULT 15
PPOV_HUMAN
```
ID   PPOV_HUMAN    STANDARD;     PRT;  1724 AA.
AC   Q9UKK3; O75903; Q9H1M6; Q14682;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Vault poly(ADP-ribose) polymerase (EC 2.4.2.30) (VPARP) (193-kDa vault
DE   protein) (PARP-related/Ialpha1-related H5/proline-rich) (PH5P).
GN   ADPRTL1 OR PARPL OR KIAA0177.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 306-319.
RX   MEDLINE=99408776; PubMed=10477748;
RA   Kickhoefer V.A., Siva A.C., Kedersha N.L., Inman E.M., Ruland C.,
RA   Streuli M., Rome L.H.;
RT   "The 193 kDa vault protein, VPARP, is a novel poly(ADP-ribose)
RT   polymerase.";
RL   J. Cell Biol. 146:917-928(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=20112770; PubMed=10644454;
RA   Still I.H., Vince P., Cowell J.K.;
RT   "Identification of a novel gene (ADPRTL1) encoding a potential
RT   poly(ADP-ribosyl)transferase protein.";
RL   Genomics 62:533-536(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Tromans A.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 94-1724 FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=96281124; PubMed=8724849;
RA   Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. V.
RT   The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 3:17-24(1996).
RN   [5]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=99198702; PubMed=10100603;
RA   Jean L., Risler J.-L., Nagase T., Coulouarn C., Nomura N.,
RA   Saler J.-P.;
RT   "The nuclear protein PH5P of the inter-alpha-inhibitor superfamily: a
RT   missing link between poly(ADP-ribose)polymerase and the
RT   inter-alpha-inhibitor family and a novel actor of DNA repair?";
RL   FEBS Lett. 446:6-8(1999).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBUNIT: BINDS TO THE MAJOR VAULT PROTEIN (MPV). THE VAULT, A
CC       LARGE RIBONUCLEOPROTEIN COMPLEX, CONTAINS THE 100-kDa MPV AND 2
```

```
Qy  478 IELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI 526
            |  || |  | :|| :|  ::|||  :  :||
Db  944 Y-IRSDG--VEIPVGETITDEHLK-SSLLYNEYIVYDVAQVNIQYLFRM 988
```

RESULT 14
YQN4_CAEEL
```
ID   YQN4_CAEEL    STANDARD;     PRT;  538 AA.
AC   Q09525;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.
GN   E02H1.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Smith A.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-
CC       RIBOSYLTRANSFERASE (EC 2.4.2.30).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z47075; CAA87379.1; -.
DR   HSSP; P26446; 1A26.
DR   WormPep; E02H1.4; CE01539.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Hypothetical protein.
SQ   SEQUENCE  538 AA;  61268 MW;  3144E25465FC7341 CRC64;
```

```
Query Match             13.9%;   Score 389;  DB 1;  Length 538;
Best Local Similarity   25.9%;   Pred. No. 4.1e-21;
Matches 148;  Conservative 86;  Mismatches 181;  Indels 156;  Gaps 24;

Qy   57 GIQVHEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLE 116
         | :|          ||||||||| :|||:||  ||| :
Db   11 GYKVH-----LCKTNIAQNNNKFYDMELLDEGGDFIVKLINGRIGYRGVTQLKDFDDLD 64

Qy  117 DAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVV------KVDS 167
         :|||:|  ||:|| : :|||||  :| ||||::|  :  |
Db   65 RAKKFFESKFYEKTHLHWEERDD-EPVPNKYAVVELATNARQTEKEVKKEEPEPEPKVDE 123

Qy  168 GPV-----RTVKPCSLDPATQNLITNIFSKEMFKNAMTL-----MNLDVKKM------ 210
         | |          | ||   | :|   :| |      : ::|  |
Db  124 KNPTRGRKKRGIVKEKKEIKEEEPVEEV--NEKLKELMKCICDEDVHLGLLKQLKFNEAF 181

Qy  211 --PLGKLTKQQIARGFEALEALEEAMKNPT---------------GDGQSLEELSS 249
         :    |::|          :                          |||:|  ::
Db  182 GRPIDCLSLAQLTTGYEILSKIEESIGGKSARRSTRGRPRVADRVLAVKSDGPSLHDINK 241

Qy  250 CFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVL--------ADIELAQTLQAAPGEEEEKVE 304
            ::|||:||| | |||  |  ::   : :  ::        ||||:| ||
Db  242 -YYSLIPHSFGFCVPPKIDSHAKIQAERELLDALKGSTEASLELKDLKKTASSKD------ 295

Qy  305 EVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLR---HVWKVN---- 357
          :      : :| | : :   | ::       |: :|    :  |:      :|  :|
Db  296 ------IYQRLYERLPCHLEPVSE----------EIAGKIGDCLAMRGPTHCYKLSLIDA 339
```

```
CC       PROTEINS OF 193 AND 240-kDa.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, BUT IS ALSO FOUND IN THE
CC       NUCLEUS, ASSOCIATED WITH MITOTIC SPINDLES.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; THE HIGHEST LEVELS ARE IN
CC       THE KIDNEY; ALSO DETECTED IN HEART, PLACENTA, LUNG, LIVER,
CC       SKELETAL MUSCLE, SPLEEN, LEUKOCYTES AND PANCREAS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC   --------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AF158255; AAD47250.1; -.
DR   EMBL; AF057160; AAC62491.1; -.
DR   EMBL; AL359763; CAC21562.1; -.
DR   EMBL; D79999; BAA11494.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR002035; vWFA.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00292; BRCT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation; Ribonucleoprotein.
FT   DOMAIN        1     94       BRCT.
FT   DOMAIN      256    569       NAD-BINDING (BY SIMILARITY).
FT   DOMAIN      876   1046       VWFA.
FT   DOMAIN     1562   1724       INTERACTION WITH THE MAJOR VAULT PROTEIN.
FT   DOMAIN       19     25       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                (POTENTIAL).
FT   DOMAIN     1237   1249       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                (POTENTIAL).
FT   CONFLICT    519    519       S -> P (IN REF. 1).
FT   CONFLICT    873    873       S -> N (IN REF. 2 AND 4).
FT   CONFLICT    897    897       Q -> E (IN REF. 2 AND 4).
FT   CONFLICT    899    899       A -> T (IN REF. 1).
FT   CONFLICT    936    936       A -> T (IN REF. 1).
FT   CONFLICT    936    936       A -> M (IN REF. 3).
FT   CONFLICT   1012   1012       V -> I (IN REF. 4).
FT   CONFLICT   1065   1065       A -> V (IN REF. 3 AND 4).
FT   CONFLICT   1080   1080       R -> L (IN REF. 3).
FT   CONFLICT   1108   1108       C -> R (IN REF. 3).
FT   CONFLICT   1265   1265       G -> A (IN REF. 2 AND 4).
FT   CONFLICT   1280   1280       G -> R (IN REF. 2 AND 4).
FT   CONFLICT   1328   1328       T -> P (IN REF. 3).
FT   CONFLICT   1331   1331       T -> A (IN REF. 3).
FT   CONFLICT   1394   1394       A -> S (IN REF. 3).
FT   CONFLICT   1459   1459       Y -> S (IN REF. 3).
FT   CONFLICT   1550   1550       P -> L (IN REF. 3).
FT   CONFLICT   1555   1555       V -> L (IN REF. 1).
FT   CONFLICT   1564   1564       T -> I (IN REF. 3).
FT   CONFLICT   1656   1656       P -> A (IN REF. 3).
SQ   SEQUENCE   1724 AA;  192587 MW;  A009F34946OEDC CRC64;

Query Match             9.9%;  Score 276.5;  DB 1;  Length 1724;
Best Local Similarity  22.1%;  Pred. No. 3.7e-12;
Matches  122;  Conservative  104;  Mismatches  211;  Indels  115;  Gaps  21;

Qy      3 PKRKAS--VQTEGSKKQRQGTEEEDSFRSTAEALRAA------PADNRVIRVDPSCPFSR 54
          |  :||   ::|||          |||:       ::         | |  :   :
Db    104 PDQKASSSEVKTEGLCPD-SATEEEDTVELTEFGMQNVEIFHLPQDFEVAKYNT---LE 158
```

```
Qy   55 NPGIQVHEDYDCTLNQTNIGNNNKFYIIQ--LLEEG---SRFFCWNRWGRVGEVGQSKM 109
        |:::          :| :: : :::   :::     ||:|
Db  159 KVGMEGGGQEAVVVELQCSRDSRDCPFLISSHFLLDDGMETRRQFAIKKTS-------- 208

Qy  110 NHFTCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGP 169
        ||| :|  :    ::  | ::    :  |   | |      || |
Db  209 ------EDASEYFENYIEELKKQGFLLREHFTPE------ATQLASEQLQALL------- 249

Qy  170 VRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQOIARGFEALEA 229
        : :| :|      :|: :: : :        |:   :: ::
Db  250 LEEVMNSSTLSQEVSDLVEMIWAEALGHLEHMLLK------PVNRISLNDVSKAEGILLL 303

Qy  230 LEEAMKNPTGDGQSLEELSSC---FYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLAD- 285
        | : ||         : |: |     || :::||    |||  :| :|
Db  304 VKAALKN---GETAEQLQKMTEFYRLIPHK-GTMPKEVNL-GLLAKKADLCQLIRDM 356

Qy  286 IELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSY 345
        : :| :   |              :| :  :::|  | :| ::| ::  :|
Db  357 VNVCETNLSKPNP--------PSLAKYRALRCKIEHVEQNTEEFLRVRKEVLQNHHSK 406

Qy  346 RCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGL---RIMPHSG- 401
        :  :::|  |  ||   |  |||  |   | :| | ::   :|      |
Db  407 SPVDVLQIFRVGRVNETTEFL--SKLGNVRPLLHGSPVQNIVGILCRGLLLPKVVEDRGV 464

Qy  402 -----GRVGKGIYFASENSKSAGYVTTMICGGHQVGYMFLGEVALGKEHHITIDDPSLKS 456
             ||||| ||:     || || :| :::     |:  |  |:  |||::|
Db  465 QRTDVGNLGSGIYFSDSLSTSIKYSHPGETDGTRL--LLICDVALGKCMDLHEKDFSLTE 522

Qy  457 PPPGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLIYKES 516
        |||    |::|    | :|| |:|  |||  ||  |        ::  |
Db  523 --------------------------------------ASVTTDFEDDEFVVYKTN 554

Qy  517 QCRLRYLLEIHL 528
        | : :::: : |
Db  555 QVKMKYIIKFSM 566
```

Search completed: August 29, 2002, 08:02:28
Job time: 305 sec

us-09-701-586b-10.rsp

.Mon·Sep 9 08:38:37 2002

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 29, 2002, 08:01:37 ; Search time 117.38 Seconds
                                              (without alignments)
                                  778.168 Million cell updates/sec

Title:           US-09-701-586B-10
Perfect score:   2789
Sequence:        1 MAPKRKASVQTEGSKKQRQG.........EYLIYKESQCRLRYLLEIHL 528

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters:        562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_19:*
                 1: sp_archea:*
                 2: sp_bacteria:*
                 3: sp_fungi:*
                 4: sp_human:*
                 5: sp_invertebrate:*
                 6: sp_mammal:*
                 7: sp_mhc:*
                 8: sp_organelle:*
                 9: sp_phage:*
                10: sp_plant:*
                11: sp_rodent:*
                12: sp_virus:*
                13: sp_vertebrate:*
                14: sp_unclassified:*
                15: sp_rvirus:*
                16: sp_bacteriap:*
                17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
                  %
Result         Query
No.   Score  Match Length DB ID        Description
----  -----  ----- ------ -- --------  -----------
 1    2785   99.9   528   11 Q91YR6    Q91yr6 mus musculu
 2    2253.5 80.8   533    4 Q96CG2    Q96cg2 homo sapien
 3     823   29.5   612    5 Q9TX06    Q9tx06 dictyosteli
 4     719.5 25.8   653   10 O50017    O50017 zea mays (m
 5     705   25.3  1014   11 Q921K2    Q921k2 mus musculu
 6     691   24.8   635   10 O81294    O81294 arabidopsis
 7     689   24.7   607   13 Q9PS82    Q9ps82 gallus gall
 8     681   24.4   607   13 Q9PS81    Q9ps81 xenopus. na
 9     624   22.4   983   10 Q9ZP54    Q9zp54 arabidopsis
10     624   22.4  1009   10 Q9SJW4    Q9sjw4 arabidopsis
11     612.5 22.0   593    5 Q9TX05    Q9tx05 drosophila
12     584   20.9   969   10 O24570    O24570 zea mays (m
13     584   20.9   980   10 Q9ZSV1    Q9zsv1 zea mays (m
14     576   20.7   945    5 Q9N4H4    Q9n4h4 caenorhabdi
15     481   17.2   727    5 Q9XUA5    Q9xua5 caenorhabdi
16     303   10.9  2276    5 Q9TXQ1    Q9txq1 caenorhabdi
17     280   10.0   815   10 Q9SWB4    Q9swb4 glycine max
18     265    9.5   815   10 Q9FK91    Q9fk91 arabidopsis
19     148    5.3  1181    5 Q9XZ37    Q9xz37 drosophila
20     148    5.1  1181    5 Q9VBP3    Q9vbp3 drosophila
21     143    5.1  1327    4 O95271    O95271 homo sapien
22     131.5   4.7   363    4 Q9H8R9    Q9h8r9 homo sapien
23     121.5   4.4   924    4 Q969W4    Q969w4 homo sapien
24     121.5   4.4  1223    4 Q9UFI5    Q9ufi5 homo sapien
25     121    4.3  1092    4 O50236    O50236 zymomonas m
26     181   12    4.3   181   12 O55721    O55721 chilo iride
27     120.5   4.3  1166    4 Q9H2K2    Q9h2k2 homo sapien
28     120.5   4.3  1265    4 Q9HAS4    Q9has4 homo sapien
29     120    4.3  5198    5 O76518    O76518 caenorhabdi
30     118    4.2   935   10 Q9CA26    Q9ca26 arabidopsis
31     117    4.2   954   10 P93826    P93826 arabidopsis
32     115.5   4.1   451    5 Q9V635    Q9v635 drosophila
33     112    4.0  1203   16 Q9CJI9    Q9cji9 lactococcus
34     111.5   4.0   261    4 Q9H8F2    Q9h8f2 homo sapien
35     111.5   4.0   359    4 Q9Y4P7    Q9y4p7 homo sapien
36     110.5   4.0  1327   11 Q61595    Q61595 mus musculu
37     108.5   3.9  1849    2 Q9S4K2    Q9s4k2 lactobacill
38     108    3.9  1412    4 Q96RT1    Q96rt1 homo sapien
39     107.5   3.9   842    4 Q9HCR1    Q9hcr1 homo sapien
40     107    3.8  1330    6 O97961    O97961 vulpes vulp
41     107    3.8  1763   11 Q9JKX5    Q9jkx5 mus musculu
42     106    3.8  1065    2 Q9AHK7    Q9ahk7 borrelia bu
43     105    3.8   849    2 Q9AHL2    Q9ahl2 borrelia bu
44     105    3.8  1065    2 Q9AHK8    Q9ahk8 borrelia bu
45     104.5   3.7  1342   10 Q9FKN5    Q9fkn5 arabidopsis
```

ALIGNMENTS

```
RESULT 1
Q91YR6
ID   Q91YR6        PRELIMINARY;        PRT;        528 AA.
AC   Q91YR6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 59.4 KDA PROTEIN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014870; AAH14870.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 528 AA; 59413 MW; 82EFB0C498EB5F74 CRC64;

Query Match             99.9%;  Score 2785;  DB 11;  Length 528;
Best Local Similarity   99.8%;  Pred. No. 6e-223;
      Matches  527;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy      1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60

Qy     61 HEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLEDAKK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 HEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLEDAKK 120

Qy    121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLD 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLD 180

Qy    181 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAAMKNPTGD 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  181 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE 300
        |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db  241 GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELEVQTLQAAPGEEE 300

Qy  301 EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWKVNREG 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWKVNREG 360

Qy  361 EGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGY 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 EGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGY 420

Qy  421 VTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIEL 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 VTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIEL 480

Qy  481 ELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
```

RESULT 2
Q96CG2
ID   Q96CG2
AC   Q96CG2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 60.1 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014260; AAH14260.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 533 AA; 60089 MW; 6296A0E439CC7767 CRC64;

```
Query Match             80.8%;  Score 2253.5;  DB 4;  Length 533;
Best Local Similarity   80.9%;  Pred. No. 9.4e-179;
Matches  432;  Conservative  37;  Mismatches  58;  Indels  7;  Gaps  4;

Qy    1 MAPKRKASVQTEG--SKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58
        ||||.|..||||    ||| || ||||.||||||||.|.||..|.|||||.||.|.|||
Db    1 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60

Qy   59 QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117
        ||.|||.||||||||.|||||||||||..:.|||||||||||||||||||.||||.||
Db   61 QVYEDYNCTLNQTNIENNNNKFYIIQLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTV--V 174
        ||||| |||.|||||.|.||||||.:|..||||||..|.:.|||||||||.||||||:||
Db  121 AKKDFEKKFPEKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 180

Qy  175 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234
        .|||||||||.||||||||||||||.| ||:||||||||||:|||||||||||||||.:.
Db  181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEAL 240

Qy  235 KNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
        |.||| |||||||||.||||||||||.| |||||||..|||||||||||||||||.||
Db  241 KGPTDGGQSLEELSSHFYTVIPHNFGHSOPPPINSPELLQAKKDMLLVLADIELAQALQA 300

Qy  295 APGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVW 354
        .  | ||.|||||||||||||||||||||||||.|||.|||||.|||||:.||.|:|.|
Db  301 V-SEQEKTVEEVPHPLDRDYQLLRCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIW 359
```

```
Qy  355 KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 414
        ||||:|| || ||||||||| ||||||||.|||||||||||||||||||||||||||||
Db  360 KVNQEEEDRPQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 419

Qy  415 SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
        |||||||  |.|| |.|||||||||||.|||| :|:|.|||||||||||||||| ||||
Db  420 SKSAGYVIGMKCGAHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP 479

Qy  475 AQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
        :||.||||||| ||||||.|| || | |||:||||||||:||||||||||::|
Db  480 TQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQSQCRLRYLLEVHL 533
```

RESULT 3
Q9TX06
ID   Q9TX06
AC   Q9TX06;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
SQ   SEQUENCE 612 AA; 69241 MW; CB340F7A88FF2364 CRC64;

```
Query Match             29.5%;  Score 823;  Pred. No. 9.9e-60;  DB 5;  Length 612;
Best Local Similarity   38.4%;
Matches  209;  Conservative  79;  Mismatches  204;  Indels  52;  Gaps  17;

Qy    6 KASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED-- 63
        .|:.| :  ||.  : :: .|:. :||.| : :       :: | : :    .:.:
Db   96 EAETKTAASDLDDSSSESEEDEKNQISVKIKGRAAN------DPHFPDSR---XHVVENGK 146

Qy   64 --YDCTLNQTNIGNNNNKFYIIQLLE--EGSRFFCWNRWGRVGEVGQSKMNHF--TCLED 117
          :| | |||.|.:.||||.|||||||    |.: || |||||..|:| |.:   .||.
Db  147 DVYDATLNQTEISQNNNKXYIIQLLEADDGSSYSVWNRWGREGLKGQSSRKDFGKGGLNQ 206

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQE--AVVKVDSGPVRTVV 174
        |  .| .||:|||||.: ::|..|..||.||:|::  :.|.:    |:.:  . .|.:
Db  207 AISLFCSKFYEKTKNTFTDRANFKKVAGKYDMIELDYSTDSKPKNGASTATTTTKKVV 266

Qy  175 ---KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALE 231
           || |||:.|:::|:.:|.:::|.:.: :.|  |.|:.|||||:|:|:|.|.:|:.:.:
Db  267 EHKKECSLDERVQELVKLIFDVKMMERTMTEAKYDLKKMPLGKLSKNQITKGYLVLKQIE 326

Qy  232 EAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQT 291
        :.:.. :   |.|| |||.||.:|||.||.| .|| |||.::::..:| :|:|||||:|
Db  327 DVMGGKS--GESLSTLSSRFYTIIPHAFGMSVPPVINTNQMLIEKMNMLQNLADIEIATN 384

Qy  292 LQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLR 351
        ::.:  :::.:    ::.  :: ::.  :  :.: : |:.: :.|:. ||| .|  .::
Db  385 IIKDSESDESNI------LELHYAKLKTDIQPLDENSCEYKNILLYVKNTYQGGKKPTIV 438

Qy  352 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKG 407
        ::.|::.:|| ||.: | .|||| ||||| .:.:.|:|:|.|||:.|    |||.:.||
Db  439 NIFKIDRDGEADRYKTKKHLGNRKLLWHGSRLTNYASIISQGLRIAPPEAPVSGYRFGKG 498
```

```
Qy   408 IYFASENSKSAGYVTTMHCGGHQVGY-MFLGEVALGKEHHITID---DPSLKSPPPGFDS 463
             ::|| |||  :: :   ||||||||:|||||      : ||: |    :   |||||||||
Db   499 VYFADCMSLSANYCRTV---GFLIDFCMLLGDVALGK----TADLARDTYMEKPQPNSHS 551

Qy   464 VIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRY 522
             : | | ||||| ||   |  ||||||  |   |: :   ||  :     |: |
Db   552 TWALGTVEPDPKVFFQ---DTESCITPYGQMI-PSQHKGVSCYEHQVVVYDVAQVHKKY 606

Qy   523 LEI 526
             |:|
Db   607 LQL 610


RESULT  4
ID   O50017
AC   O50017;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babyichuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants possess two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ222588; CAA10888.1; -.
     HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
SQ   SEQUENCE 653 AA; 72995 MW; 5FD01923C4ABCD1D CRC64;

Query Match             25.8%;  Score 719.5;  DB 10;  Pred. No. 4.4e-51;
Best Local Similarity   35.3%;
Matches  190;  Conservative   91;  Mismatches  205;  Indels   53;  Gaps   17;

Qy    13 GSKKQRQGTEEDSFRS----TAEALRAAPADNRV---IRVDPSCFFSRNPGIQVHED-Y 64
             | |:: |  ||| ||:     ||||||||||:    ||| | |:  ||| ||   |
Db   138 GAKEVIKGGDEEVEKKEKMVTATKKGAAVLDOHIPDHIKVNHV------LQVGDEIY 190

Qy    65 DCTLNQTNIGNNNNKFYIIQLLEE--GSRFFCWNRWGRVGEVGQOSKMNHFTCLED-AKKD 121
             | |||||:| ||||||||||:|   | |:  ||||||||| | |:|  ||  | ||
Db   191 DATLNQTNVGDNNNKFYIIQVLESDAGGSFMYNRWGRVGRGVGRGQDKLHGPSPTRDQAIYE 250

Qy   122 FKKKFWEKTKNKWEEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKVDSGPVRTVKPCSLD 180
             | ||| : ||:|  :|  | |  :|||  |:    ||     |:
Db   251 FEGKFHNKTNNHWSDRKNFKCYAKKYTWLEMOYGETEKE-----IEKGSITDQIKETKLE 305

Qy   181 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 240
             ||:| | :|| |:||    : : ||:| ||||||: | | |
Db   306 TRIAQFISLICNISMMKQRMVEIGYNAEKLPLGKLRATILKGYHVLKRISDVISK--AD 363

Qy   241 GQSLEELSSCFYTVIPHNFG--RSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGE 298
             :: ||| | |||||||:|||    | | |:| : |  |:|||:| |: | | :
Db   364 RRHLEQLTGEFYTVIPHDFGFRKMREFIIDIPQKLKAKLEMVEALGEIEIATKLL----- 418

Qy   299 EEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRCPNLRHVWK 355
             :: ||| |  | || ||:| ||  |   :  ||  ||:||:   : :| : ::|:
Db   419 -EDDSSDQDDPLYARYKQLHCDFTPLEADSDEYSMIKSYLRNTHGKTHSGYTVDIVQIFK 477


Qy   356 VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFA 411
             |::  |  || ::||||||:||||| : :  || :|| ||: |      ||||||:||
Db   478 VSRHGETERFQKFASTRNRMLLWHGSRLSNWAGIISQGLRIAPPEAPVTGVMFGKGVYFA 537

Qy   412 SENSKSAGYVTTMHCGGHQV---GYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 468
             |   ||:   |   |     |:|| ||:||||||| :  :||: |||| |:|||
Db   538 DMFSKSSANY-----CYASEACRSGVLLCEVALGDMNELLNADYDANNLPKGKLRSKGVG 592

Qy   469 QTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEIH 527
             ||: |  :: :    ||||| || :  ||  ||  :|||||:| :: :| :::::|: |
Db   593 QTAPNM---VESKVADDGVVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMRVVLHVN 647


RESULT  5
ID   Q921K2
AC   Q921K2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE)
DE   POLYMERASE).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC012041; AAH12041.1; -.
KW   Transferase.
SQ   SEQUENCE 1014 AA; 112721 MW; 5DCE68E4CB3F46EB CRC64;

Query Match             25.3%;  Score 705;  DB 11;  Length 1014;
Best Local Similarity   33.6%;  Pred. No. 1.4e-49;
Matches  182;  Conservative  101;  Mismatches  214;  Indels   44;  Gaps   17;

Qy     1 MAPRKKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
             :|| |: :   |: ||| :|   |:     || : |  :| : |:
Db   494 VAPKGKSAAPSKKSK---GAVKEEGVNKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 549

Qy    61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
             ::  | || :  ::  |:||:: ||: |   ||::  ||||||| | |
Db   550 GKVFSATLGLVDIVKGTNSYKLQLLEDDKESRYWIFRSWGRVGTVIGSNKLEQMPSKED 609

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKP- 176
             | : | |:: |:| | | :: | |: |: :: |:| | :     |:||||| |:: ||
Db   610 AVEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL-----TVKPG 660

Qy   177 --CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234
               | | || :: : :|:| |   | ::  :|:|||||||: :: | |      |: |
Db   661 TKSKLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQAV 720

Qy   235 KNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
             |:  ::: | ::| |||| |||:|||  ::|: | || |  |:| |:|| | |:
Db   721 SQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLR 779

Qy   295 APGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ--         350
                 :: : :  |  | | : :::|  : :::|::| |::|:
Db   780 APGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ-----TGNSYRCPNL 833

Qy   351 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGK 406
             || : ||:  |  || ::||||||:||||| : :  || :|| ||: |      ||||
Db   834 IDIFKIERRGESQRYKPFQLHNRRLLWHGSRTNFAGIISQGLRIAPPEAPVTGYMFGK 893

Qy   407 GIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIA 466
             |||||   ||:|   |   |  ||||| ||:||||||: :  ::|: ||||||:|||
Db   894 GIYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSVKG 950

Qy   467 RGQTEPDPAQDIELELDGQPVVPQGPPVVQCPS-FKSSSFSQSEYLIYKESQCRLRYLLE 525
             :| ||: |  :: :    ||||| || :  || :||  :|||||:| :: :| :::::|:
```

```
Db  951 LGKTPDPSASITLE---GVEVPLGTGI--PSGVNDTCLLYNEYIVYDIAQVNLKYLLK 1004
Qy  526 I 526
Db 1005 L 1005
```

```
RESULT  6
081294
ID   081294    PRELIMINARY;    PRT;   635 AA.
AC   081294;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE).
GN   T14P8.19 OR AT4G02390.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   WASHU;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RA   Kalicki J., Elliott G., Cloud J.;
RT   "The sequence of A. thaliana T14P8.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Waterston R.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF069298; AAC19283.1; -.
DR   EMBL; AL161494; CAB80732.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
KW   Transferase.
SQ   SEQUENCE   635 AA;  72017 MW;  E3F1CBE4D367A377 CRC64;
```

```
Query Match             24.8%;   Score 691;  DB 10;  Length 635;
Best Local Similarity   34.6%;   Pred. No. 9.9e-49;
Matches  189;  Conservative  91;  Mismatches 220;  Indels 220;  Gaps  46;
```

```
Qy    2 APKRRASVQTEGSKKQRGTEEEDSFRS--TAEALRAAPADNRVIRVDPSCPFSRNPGIQ 59
         || :: :    |   |:| |||       :| ||   ::     ::  |:
Db  109 APVKSSNDEAE--DDNNGFEEEKKEEKIVTATKKGAAVLDQWI----PDEIKSQYHVLQ 161
Qy   60 VHED-YDCTLNQTNIGNNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMN-HFTCL 115
         |  : |  ||||||: :|||||| ::|:||  :  || |||||||| :||||::  : |
Db  162 RGDDVYDAILNQTNVRDNNNKFFVLQVLESDSKKTMVYTRWGRVGVKGQSKLDGPYDSW 221
```

```
Qy  116 EDAKKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVK 175
         : ::   ||  :||:|||  :: ::|:   | : :| :    ::  : || :||
Db  222 DRAIEIFTNKFNDKTKNYWSDRKEFIPHKSYTWLEMDYGKEENDSPVNNDIPSSSSEVK 281
Qy  176 P--CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEA 233
           ||  | : :|    |   |  |    :   :  || | :  | :: : :|||| : |
Db  282 PEQSKLDTRVAKFISLICNVSMMAQHMEIGYNANKLPFGKISKSTISKGYEVLKRISEV 341
Qy  234 MKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPP--INSPDVLQAKKDMLLVLADIELA-Q 290
         |:  | |  |: || : : |||:|||:||||||    |||:|::| |  |:: |:| ||| :
Db  342 IDR--YDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATK 399
Qy  291 TLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRC 347
         || |   ||| :    ::    ||  | :: |   :  : :|| :  ||:| | | ::|
Db  400 ------PLYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYT 452
Qy  348 PNLRHVKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGR 403
         ||| | | | ::|  |  : :||:|| |||:|||:|| | |||::|||:|  |   ::||
Db  453 VEIAQLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYM 512
Qy  404 VGKGIYFASENSKSAGYVTTMHC---GHQVGYMFLGEVALGKEHHTIDDPSLKSPPPG 460
         |||:||| ||:|||| | |||    | :|  | ::|| :||| | |:| |::  ||| :
Db  513 FGKGVYFADMFSKSANY----CYANTGANDGVLLLCEVALGDMNELLYSDYNADNLPPG 567
Qy  461 FDSVIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRL 520
         |||: | || | :|:     |  ::|  | |||| |:    || : |:|:  ||: :|:|
Db  568 KLSTKGVGKTAPNPSEAQTLE-DG--VVPLGKPVE-RSCSKGMLLYNEYIVYNVEQIKM 623
Qy  521 RYLLEI 526
         || :::
Db  624 RYVIQV 629
```

```
RESULT  7
Q9PS82
ID   Q9PS82    PRELIMINARY;    PRT;   607 AA.
AC   Q9PS82;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   607 AA;  68033 MW;  75F6EE1D30D8F402 CRC64;
```

```
Query Match             24.7%;   Score 689;  DB 13;  Length 607;
Best Local Similarity   33.9%;   Pred. No. 1.4e-48;
Matches 184;  Conservative  97;  Mismatches 205;  Indels 56;  Gaps 63;
```

```
Qy    5 RKASVQTEGSKKQRQG-TEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED 63
         ||: |: |:: |::||  |:||          ||  | :    ::  |:
Db   93 KPANMKSAGKVKEEQGPSKSEKKMKLTVKGGAAVDPDSGL--EDSAHVFEKGGKI----- 145
```

```
                                                    Matches    178;   Conservative    95;   Mismatches   215;   Indels    42;   Gaps    18;

Qy      10  QTEGSKKQRQGT-EEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHEDYDCTL    68
                :: :|||  ||:    |||  :||:     |: ::| |
Db      97  KSSGKVKEEKGSNKSEKKMKLTVKGGAIDPDS---ELEDSCHVLETGG----KIFSATL   149

Qy      69  NQTNIGNNNNKFYIQLLE--EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKKDFKKK   125
                :: |  :| :|||:         |||||:|| :  ||| | : ||||| || :|
Db     150  GLVDITRGTNSYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKEDAIEHFLNL   209

Qy     126  FWEKTKNWEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKVDSGPVRTVVKPCSLDPATQ   184
                :::| || | :   ||||||::|:|||:|:::: :
Db     210  YQDKTGNAWHS-PNFTKYPKKFYPLEIDYGQEEDVVKKLSVGAGTKSKLAKP------VQ   262

Qy     185  NLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSL   244
                 ||  | | || : :|| :: |||||:|||||||  ||||  ||| :|
Db     263  ELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSLLSQVQQAVSESLSEARLL   322

Qy     245  EELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVE   304
                 :|:|| |:| ||    ||:|| :||| ||:|| |||:||  |: ||: || :|
Db     323  -DLSNQFYTLIPHDFGMKKPPLLNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDGEK--   379

Qy     305  EVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLK----QTGNSYRCPNLRHVWKVNREG   360
                |: ||:|||:    | |||||||||  ::|| :  |||| |  |||:| | | :|
Db     380  --DPIDVKYEKIKTDIKVVAKDSEESRIICDVKNTHADTHNAYDLEVL-EIFKIDREG   435

Qy     361  KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH------SGGRVGKGIYFASENSK   416
                | ||| :| :|  ||:||| |||| || |:|| |||:||  |||:         ||| |||:| ||||
Db     436  EYQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSK   495

Qy     417  IARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLL   476
                | ||| |:  ::  ||:| |||||||||||||||| |||||||:|||||||:||||||||
Db     496  SANYCHAM--PGSPIGLILLGEVALGNMHELKAASQITKL-PKGKHSVKGLGRTAPDPSA   552

Qy     477  DIELELDGQPVVVQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI   526
                :: :: :
Db     553  TV--QLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLKYLLKL   597


RESULT   9
Q9ZP54
ID   Q9ZP54;                   PRELIMINARY;            PRT;    983 AA.
AC   Q9ZP54;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).
GN   PARP-1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LANDSBERG ERECTA;
RA   Doucet-chabeaud G., Kazmaier M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ131705; CAA10482.1; -.
DR   HSSP; P26446; 1A26.
KW   Transferase; Glycosyltransferase; NAD.
FT   CHAIN         2     983       POLY(ADP-RIBOSE) POLYMERASE.
SQ   SEQUENCE   983 AA;  111232 MW;  468E12A8EF1B6F4F CRC64;

Query Match             22.4%;   Score 624;  DB 10;  Length 983;
Best Local Similarity   33.0%;   Pred. No. 7.2e-43;
                 Matches  179;  Conservative  87;  Mismatches  209;  Indels  68;  Gaps  21;

Qy      15  KKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH----ED----YDC   66
                || |||  :| |   :  ::   | : | ::        ||:|||||        |:
Db     472  KKQRKLPFDKYKIEDTSESLVTVVKGRSAVHEAS------GLQEHCHILEDGNSIYNT   524
```

```
Qy      64  YDCTLNQTNIGNNNNKFYIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKK   120
                :   |: |||| |||||:|||||      |||||:||||||| ||||||| ||||||
Db     146  FSATLGLVDIVKGTNSYKLQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAVE   205

Qy     121  DFKKKFWEKTKNWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVKPCS   178
                | || |  ::   | |   |||||||:|:  | |: ::: |  | :
Db     206  HFLNLYEEKTGNSWHSKN-FTKYPKKFYPLEID-YGQDEEAVRKLTVSAGTKSKLAKP--   261

Qy     179  LDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLFKQQIARGFEALEALEEAMKNPT   238
                :::    ::|||  :|||| |||||||:|||| | ::  |:||||| || :| :
Db     262  ----IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILNEVQQAVSDGG   317

Qy     239  GDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGE   298
                |  |||  :| |:| |||| |||:||:|| :||||:|| |||:||  | |:| |
Db     318  SESQIL-DLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRGGNE   376

Qy     299  EEEKVEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYRCPNLRHVW   354
                |: |:   :| |||:  :| | |||||||||  ::|: ||      ||:|| | | : |
Db     377  DGDK-----DPIDINYEKLRTDIKVVDKDSEEAKIIKQYVKNTHAATHNAYDL-KVEIF   430

Qy     355  KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYF   410
                | :|||||| :|| ||:||| |||||  |:|  |||:|||  |||       ||| |||||
Db     431  RIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYF   490

Qy     411  ASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEH------HITDDPSLKSPPPGFDSV   464
                | || ::  |: |||  |||:| |:||| ||| :|          |: |||:|||:| |
Db     491  RDMVSKSANYCHTSQ-ADPIGLILLGEVALGNMYELKNASHIT-----KLPKGKHSV   541

Qy     465  IARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLL   524
                | |::   :  :  ||: | |||||||||||||||||||||||:|||||||:|||||||
Db     542  KGLGKTAPDPT--ATTTLDG--VEVPLGNGIS-TGINDTCLLYNEYIVYDVAQVNLKYLL   596

Qy     525  EI   526
                ::
Db     597  KL   598


RESULT   8
Q9PS81
ID   Q9PS81;                   PRELIMINARY;            PRT;    607 AA.
AC   Q9PS81;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Xenopus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae.
OX   NCBI_TaxID=8353;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   607 AA;  67496 MW;  54CDEBBE22079886 CRC64;

Query Match             24.4%;   Score 681;  DB 13;  Length 607;
Best Local Similarity   33.6%;   Pred. No. 6.3e-48;
```

```
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE   1009 AA;  114133 MW;  CDE6E41CC2A3A2DB CRC64;

Query Match             22.4%;  Score 624;  DB 10;  Length 1009;
  Best Local Similarity   33.0%;  Pred. No. 7.5e-43;
  Matches  179;  Conservative   87;  Mismatches 209;  Indels  68;  Gaps  21;

Qy         15 KKQRGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH----YDC   66
              ||||| :: : :: |::::| || :  |:||  :  : :  | ---- :
Db        498 KKQRKLPFDKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT  550

Qy         67 TLNQTNIGNNNKFYIIQLLEE--GSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDF  122
              || :: :: :: |::||| ||  | |||| ||  |||||  ||:||||| | :: :
Db        551 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGGNKVEEMS-KSDAVHEF  609

Qy        123 KKKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSL  179
              |: |||||||   |::|| :|||  :||:||   | :  ||  | :| |
Db        610 KRLFLEKTGNTWESWEQKTNFQKQPGKFLPLDIDYGVNKQVA----KKEPFQT---SSNL  662

Qy        180 DPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQOIARGFEAL---EALEEAMK  235
              |  |::|  :  | | :  ::|::::|  :|||| | || |:  |:||   | | |
Db        663 APSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQRLLTESDP  722

Qy        236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA  295
              || |:|| :||:: ::| | |  |    ::|| || :  |  :  |::||  | ||
Db        723 QPTMKESLLVDASNRFFTMIP----SIHPHIIRDEDDFKSKVKMLEALQDIEIASRI---  775

Qy        296 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------  348
              ||::::|  :|||| :: : |  | |  : |:|:|| |  |   |:| ::|
Db        776 VGFDVDSTES----LDDKYKKLHCDISPLPHDSEDYRLIEKYL---NTTHAPTHTEWSL  827

Qy        349 NLRHVWKVNREGEGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRI---MPHSGGR  403
              ||| | |::|| | |:| ||   || |  ||||: |||||  |: |:|   | || ||
Db        828 ELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAPATGYM  887

Qy        404 VGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDS  463
              || ||||||:  ||||: || ||| | | |:||||||: |:|| ::||  | |||| |
Db        888 FGKGIYFADLVSKSAQYCYT--CKKNPVGLMLLSEVALGEIHELT-KAKYMDKPPRGKHS  944

Qy        464 VIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYL  523
              | | ::  :|  ||:|||:|| |  | | || |||  ::|| :|::|:| |||| ||
Db        945 TKGLGKKVP----QDSEFAKWRGDVTVPCGKPVS-SKVKASELMYNEYIVDTAQVKLQFL 1000

Qy        524 LEI 526
              |::
Db       1001 LKV 1003

RESULT 10
Q9SJW4
ID   Q9SJW4                PRELIMINARY;       PRT;   1009 AA.
AC   Q9SJW4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE POLY (ADP-RIBOSE) POLYMERASE.
GN   AT2G31320.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. COLUMBIA;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006593; AAD20677.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
```

RESULT 11
Q9TX05
```
ID   Q9TX05                PRELIMINARY;       PRT;   593 AA.
AC   Q9TX05;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Drosophila sp. (fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7242;
RN   [1]
```

RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   593 AA;   68018 MW;   D9BA37E3B8E7CCD CRC64;

Query Match
Best Local Similarity   22.0%;   Score 612.5;   DB 5;   Length 593;
                        Pred. 31.4%;   Pred. No. 3e-42;
Matches 166;   Conservative 102;   Mismatches 222;   Indels 39;   Gaps 17;

Qy   16 KQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED----YDCT 67
Db   80 KSRIPEKETTKSLNSNSIYTKSMPV-SRTFKVKDGLAVDPDSGLEDIAHVYDSNNKYSVV 138

Qy   68 LNQTNIGNNNKFYIIQLL--EEGSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKK 124
Db   139 LGLTDIQRNKNSYYKVQLLKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKE 198

Qy   125 KFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLDPATQ 184
Db   199 IYADKTGNEYEQRDNFVKRTGRMYPIEIQ--YDDDQKLVKHESHFFTS-----KLEISVQ 251

Qy   185 NLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEEAMKNPTGDGQSL 244
Db   252 NLIKLLFDIDSMNKTLMEFHIDMDKMPLGKLSAHQIQSAYRVVKEIYNVLECGSNTAK-L 310

Qy   245 EELSCSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVE 304
Db   311 IDATNRFYTLIPHNFGVQLPTLIETHQQIEDLRQMLDSLAEIEVAYSI-----IKSEDVS 365

Qy   305 EVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKYNREGE 361
Db   366 DACNPLDNHYAQIKTQLVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKVSRQGE 425

Qy   362 GDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRI----MPHSGGRVGKGIYFASENSKS 417
Db   426 ARREKPFKKLHNRKLLWHGSRLTNFVGILSHGLRIAPPEAPVTGYMFGKGIYFADMVSKS 485

Qy   418 AGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQD 477
Db   486 ANYCCTSQ--QNSTGLMLLSEVALGDMMECT-SAKYINKLSNNKHSCFGRGRTMPDPTKS 542

Qy   478 IELELDGQPVVPQGGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI 526
Db   543 Y-IRSDG--VEIPYGETITDEHLK-SSLLYNEYIVYDVAQVNIQYLFRM 587

RESULT 12
ID   024570   PRELIMINARY;   PRT;   969 AA.
AC   024570;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.

OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babiychuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants possess two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ222589; CAA10889.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE   969 AA;   109128 MW;   EB23AC62EEC14009 CRC64;

Query Match
Best Local Similarity   20.9%;   Score 584;   DB 10;   Length 969;
                        Pred. 30.9%;   Pred. No. 1.5e-39;
Matches 164;   Conservative 94;   Mismatches 201;   Indels 72;   Gaps

Qy   33 ALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED---YDCTLNQTNIGNNNKFYIIQ 84
Db   468 ALESSKGSTVTVKVKGRSAVHESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQ 527

Qy   85 LLE--EGSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKKFWEKTKNKWEE---R 137
Db   528 IIEQDDGSECYFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFEKTGNSWEAWECK 586

Qy   138 DRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVK--PCSLDPATQNLITNIFSKEM 195
Db   587 TNFRKQPGRFYPLDVD--------YGVKKAPKRKDISEMKSSLAPQLLELMKMLFNVET 637

Qy   196 FKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEAEAMKNPTGDGQSLEE----LSSC 250
Db   638 YRAAAMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAASNR 697

Qy   251 FYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVEEVPHPL 310
Db   698 FFTLIP---SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDES------L 746

Qy   311 DRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP-------NLRHVWKVNREGEGD 363
Db   747 DDKYMKLHCDITPLAHDSEDYKLIEQYLLNT---HAPTHKDWSLELEEVFSLDRDGELN 802

Qy   364 RFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSKSA 418
Db   803 KYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVSKSA 862

Qy   419 GYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPA 475
Db   863 QY------CYVDRNNPVGLMLLSEVALGDMYELK-KATSMDKPPRGKHSTKGLGKTVPLES 916

Qy   476 QDIELELDGOPVVPOGGPPVOGPSFKSSSSFSQSEYLIYKESQCRLRYLLEI 526
Db   917 EFVKWRDD---VVVPCGKPVP-SSIRSSELMYNEYIVYNTSQVKMQFLLKV 963

RESULT 13
ID   Q9ZSV1   PRELIMINARY;   PRT;   980 AA.
AC   Q9ZSV1;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```
DE  POLY(ADP)-RIBOSE POLYMERASE (EC 2.4.2.30).
GN  PARP1.
OS  Zea mays (Maize).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC  Panicoideae; Andropogoneae; Zea.
OX  NCBI_TaxID=4577;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99026291; PubMed=9808734;
RA  Mahajan P.B., Zuo Z.;
RT  "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL  Plant Physiol. 118:895-905(1998).
DR  EMBL; AF093627; AAC79704.1; -.
    HSSP; P26446; 1A26.
DR  InterPro; IPR001357; BRCT.
DR  InterPro; IPR001290; PARP.
DR  InterPro; IPR004102; PARP_reg.
DR  InterPro; IPR003034; SAP.
DR  InterPro; IPR001510; Znf-PARP.
DR  Pfam; PF00533; BRCT; 1.
DR  Pfam; PF00644; PARP; 1.
DR  Pfam; PF02877; PARP_reg; 1.
DR  Pfam; PF00645; zf-PARP; 2.
DR  ProDom; PD004675; Znf-PARP; 2.
DR  SMART; SM00292; BRCT; 1.
DR  PROSITE; PS50172; BRCT; 1.
DR  PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW  Transferase; Glycosyltransferase.
SQ  SEQUENCE   980 AA;  110475 MW;  9D8AED26BC37E5C1 CRC64;

Query Match          20.9%;    Score 584;  DB 10;  Length 980;
Best Local Similarity 30.9%;   Pred. No. 1.5e-39;
Matches   164;  Conservative  94;  Mismatches  201;  Indels  72;  Gaps  20;

Qy   33 ALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED----YDCTLNQTNIGNNNNKFYIIQ 84
            ||   |           |:|:|          :: ||          |:|| :: :::|
Db  479 ALESSKGSTVTVKVKGRSAVHESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQ 538

Qy   85 LLE--EGSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDFKKKFWEKTKNKWEE---R 137
         :||   |||||: :|||||||    |||||||:|||||||||||||||:||||     |
Db  539 IIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEKTGNSWEAWECK 597

Qy  138 DRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVK-PCSLDPATQNLITNIFSKEM 195
         ||||||||||||||||||||||||||||||||||||||  | | |  | || :| :|:|
Db  598 TNFRKQPGFPYPLDVD-------YGVKKAPKRKDISEMKSSLAPQLLELMKMLFNVET 648

Qy  196 FKNAMTLMNLDVKKMPLGKJTKQQIARGFEALEALEEAMKNPTGDGQSLEE-----LSSC 250
         ::  |  ::||||||||:|  |  || :||||||| |          |:        |:
Db  649 YRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAASNR 708

Qy  251 FYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVEEVPHPL 310
         |: ||||||  :        : |  | | ||||| || |    |   ::
Db  709 FFTLIP----SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDES------L 757

Qy  311 DRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP-------NLRHVWKVNREGEGD 363
         |  | :|:| : || | |:|||:|  ||:  |          |    | | :|:|
Db  758 DDKYMKLHCDITPLAHDSEDYKLIEQYLLNT----HAPTHKDWSLELEEVFSLDRDGELN 813

Qy  364 RFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRIMPH-----SGGRVGKGIYFASENSKSA 418
          :|   ::|  ||||||| |  : |:|: ||||:         |||:||:||| | |||
Db  814 KYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVSKSA 873

Qy  419 GYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPA 475
         | :|         | :|  : :|| |        |    :        : |||
Db  874 QY------CVVDRNNPVGLMLLSEVALGDMYELK-KATSMDKPPRGKHSTKGLGKTVPLES 927

Qy  476 QDIELELDGQPVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEI 526
         :       |   |  | |      :||::  | | :||  ::: |  ||:|
Db  928 EFVKWRDD---VVVPCGKPVP-SSIRSSELMYNEYIVVNTSQVKMQFLLKV 974
```

```
RESULT  14
Q9N4H4
ID   Q9N4H4        PRELIMINARY;    PRT;    945 AA.
AC   Q9N4H4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 108.0 KDA PROTEIN.
GN   Y71F9AL.18.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw-Cordum H., Scott K., Graves T.;
RT   "The sequence of C. elegans cosmid Y71F9AL.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024200; AAF36011.1; -.
     HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   945 AA;  108006 MW;  1D0A62C954BC6AD9 CRC64;

Query Match          20.7%;    Score 576;  DB 5;  Length 945;
Best Local Similarity 31.8%;   Pred. No. 6.6e-39;
Matches   155;  Conservative  83;  Mismatches  202;  Indels  48;  Gaps  14;

Qy   64 YDCTLNQTNIGNNNNKFYIIQLLEEGSR--FFCWNRWGRVG-EVGQSKMNHFTCLEDAKK 120
         | :||   | |:|: | :||||| :: | :||| :|||||| |||::|:||   | |
Db  477 YQATLSFTDLFQNKNSYYKIQLLKDDQRENYVFRSWGRVGTEVGGNKHESYSNSNEAIL 536

Qy  121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEASQEAVVKVDSGPVRTVVKPCSLD 180
         :|: :|:|||||:| :| |  : :  |: |:|  :|:  | | ::  |  |:: |: |
Db  537 KFQDVFHEKTKNDWIYRKHFRKMPGMFSYVETDYSEFAQITDTEITPGS-KTL----LP 590

Qy  181 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAEAMKNPTGD 240
              :::|:|   ::  : |   :|::||||:|   |::  | ::  ::    ::|:
Db  591 KSVKEVMSIFDVENMKSALKSFEMDVNKMPLGRLSHNQINLAFEVLNDISDLLVKLPID 650

Qy  241 GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE 300
          : :  :|| ||::||||||  | | |:| : : |:|:   :    :    : |
Db  651 ASRILDFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNMLNALLDIKFAYD-QISGGDVP 709

Qy  301 EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCP-NLRHVWKVNR 358
         ::  :   :|   ::|:|::  ||  | :||  |: |||:| | :     :||  : :||
Db  710 ASTSLGIDPVDINYQLKLCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVELIDILKLNR 769

Qy  359 EGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASEN 414
         ::|   ||:   :|||||||||:   : | ||: ||||:        |  :||| :||  |
Db  770 DNESSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPPEAPVSGYMFGKGVFFADMF 827
```

```
Qy  415 SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFDSVIARGQ---- 469
Db  828 SKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSVQGLGRQCPR 884

Qy  470 -----TEPD----PAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRL 520
Db  885 EIGSYNHPDGYTIPLGLTYMQLQGKQDV-----------DYHLLYNEFIVYDVDQIQL 931

Qy  521 RYLLEIHL 528
Db  932 KYLVRVKM 939
```

```
RESULT 15
Q9XUA5
ID   Q9XUA5      PRELIMINARY;       PRT;    727 AA.
AC   Q9XUA5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC8.1 PROTEIN.
GN   AC8.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McMurray A.A.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C. elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z83097; CAB05448.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
SQ   SEQUENCE   727 AA;  82884 MW;  530ABA8E991FFEED CRC64;
```

```
Query Match             17.2%;      Score 481;   DB 5;     Length 727;
Best Local Similarity   30.0%;      Pred. No. 3.5e;31;
Matches  133;   Conservative  68;   Mismatches 165;   Indels  78;   Gaps  13;

Qy  126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLDPATQN 185
Db  315 FHEKTKNDWIYRKHFRKMPGMFSYVET----DYSEFVGTNNGHKKKITPGSKITPGSKT 369

Qy  186 L------ITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEEAMKNP 237
Db  370 LLPKSVKEVVMSIFDVENMKSALKSFEIDVNKMPLGRLSHNQINLAFEVLNDISDLLVKL 429

Qy  238 TGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPG 297
Db  430 PIDASKILDFSNKFYTIIPHNFGMVRVPEPIDSFHKIKEKNNMLNALLDIKFAYD-QISGG 488

Qy  298 EEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCP-NLRHVWK 355
Db  489 DVPASTSLSIDPVDINYRKLKCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVELIDILK 548

Qy  356 VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFA 411
Db  549 VNRDNESSKFKRH--IGNRRLLWHGSGKMNFAGIILGQGLRIAPPBAPVSCYMFGKGVYFA 606
```

```
Qy  412 SENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFDSVIARGQT 470
Db  607 DMFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSV----- 657

Qy  471 EPDPAQDIELELDGQPVVVPQGPPVQCP----SFKSS-------SFSQ------- 507
Db  658 ------------------QGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQGKQNVDYH 697

Qy  508 ---SEYLIYKESQCRLRYLLEIHL 528
Db  698 LLYNEFIVYDVDQIQLKYLVRVKM 721
```

```
                    GenCore version 4.5
            Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      August 29, 2002, 07:57:19 ;  Search time 124.84 Seconds
                                            (without alignments)
                                            469.777 Million cell updates/sec

Title:           US-09-701-586b-10
Perfect score:   2789
Sequence:        1 MAPKRKASVQTEGSKKQRQG.........EVLIYKESQCRLRYLLEIHL 528

Scoring table:   BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:        747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters:       747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  A_Geneseq_032802:*
         1:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1980.DAT:*
         2:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1981.DAT:*
         3:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1982.DAT:*
         4:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1983.DAT:*
         5:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1984.DAT:*
         6:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1985.DAT:*
         7:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1986.DAT:*
         8:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1987.DAT:*
         9:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1988.DAT:*
        10:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1989.DAT:*
        11:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1990.DAT:*
        12:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1991.DAT:*
        13:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1992.DAT:*
        14:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1993.DAT:*
        15:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1994.DAT:*
        16:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1995.DAT:*
        17:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1996.DAT:*
        18:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1997.DAT:*
        19:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1998.DAT:*
        20:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA1999.DAT:*
        21:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA2000.DAT:*
        22:  /SIDS1/gcgdata/hold-geneseq/geneseq-embl/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	2789	100.0	528	21	AAY51178	Murine PARP1 (shor
2	2776.5	99.6	533	21	AAY51177	Murine PARP1 (long
3	2253.5	80.8	533	21	AAY51175	Human brain PARP3
4	2253.5	80.8	540	21	AAY51176	Human uterus type
5	2241.5	80.4	533	22	AAU29021	Human PARP-3 prote
6	719.5	25.8	653	21	AAY68834	A poly(ADP-ribose)
7	707.5	25.4	1013	17	AAR99642	poly(ADP-ribose) p
8	707	25.3	522	22	AAU29022	Mouse PARP-2 prote
9	702.5	25.2	1014	21	AAY58043	Human poly (ADP-ri
10	702.5	25.2	1014	21	AAU29019	Human PARP-1 prote
11	702.5	25.2	1014	22	AAB66296	Human tankyrase2 r
12		701.5	25.2	1014	20	AAY33699	Human poly(ADP-rib
13		700.5	25.1	521	22	AAB60693	Human poly(ADP-rib
14		700.5	25.1	570	22	AAY51174	Human brain PARP2
15		700.5	25.1	570	22	AAB11480	Human brain poly-A
16		699	25.1	534	21	AAB42909	Human ORFX ORF2673
17		699	25.1	534	22	AAU29023	Human PARP-2 prote
18		699	25.1	583	22	AAB47029	hPARP2. Homo sapi
19		698.5	25.0	1014	21	AAY49939	Human nuclear NAD+
20		698	25.0	534	22	AAU29020	Human PARP-2 prote
21		697.5	25.0	637	21	AAY68835	The poly(ADP-ribos
22		649	23.3	531	22	AAB93513	Human protein sequ
23		621.5	22.3	557	22	ABB66431	Drosophila melanog
24		609	21.8	1063	22	AAB47032	Fusion protein PAR
25		584	20.9	969	21	AAY68833	A poly(ADP-ribose)
26		584	20.9	980	21	AAY68839	A poly(ADP-ribose)
27		490.5	17.6	982	20	AAY28464	Maize poly ADP-rib
28		456.5	16.4	379	22	AAU21687	Novel human neopla
29		437.5	15.7	360	22	AAB47030	N-terminal fragmen
30		405.5	14.5	1010	21	AAY68840	Fusion protein of
31		378.5	13.6	294	22	AAU20129	Human DNA repair a
32		378.5	13.6	294	22	AAU21810	Novel human neopla
33		364.5	13.1	287	22	AAB47031	C-terminal fragmen
34		276.5	9.9	1730	22	AAU33242	Novel human secret
35		272	9.8	1724	21	AAY54373	cDNA sequence enco
36		272	9.8	1724	22	AAB51022	Human minor vault
37		267.5	9.6	1099	22	AAB66301	Human tankyrase2 e
38		246.5	8.8	227	22	AAU20130	Human DNA repair a
39		246.5	8.8	227	22	AAU21811	Novel human neopla
40		203.5	7.3	190	22	AAU21688	Novel human neopla
41		148	5.3	1181	22	ABB60894	Drosophila melanog
42		148	5.3	1181	22	AAB66297	Drosophila tankyra
43		143	5.1	1327	22	AAB27212	Human tankyrase I
44		143	5.1	1327	22	AAY44402	Human tankyrase.
45		143	5.1	1327	22	AAB66279	Human tankyrase1 s

ALIGNMENTS

```
RESULT 1
AAY51178
ID   AAY51178 standard; Protein; 528 AA.
XX
AC   AAY51178;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Murine PARP1 (short) homologue protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP1.
XX
OS   Mus sp.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;  99WO-EP03889.
XX
PR   05-JUN-1998;  98DE-1025213.
PR   01-MAR-1999;  99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44291.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
```

PT   conditions -
XX
PS   Claim 4; Page 71-73; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC   detection of PARP homologues and antibodies are useful for analytic
CC   binding partners, as well as for identifying PARP effectors or
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   murine PARP1 protein used in the method of the invention.
XX
SQ   Sequence 528 AA;

Query Match   100.0%; Score 2789; DB 21; Length 528;
Best Local Similarity   100.0%; Pred. No. 4.3e-239;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 mapkrkasvqtegskkqrqgteeedsfrstaealraapadnrvirvdpscpfsrnpgiqv 60

Qy   61 HEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLEDAKK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 hedydctlnqtnignnnkfyiiqlleegsrffcwnrwgrvgevgqgskmnhftcledakk 120

Qy  121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLD 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 dfkkkfwektknkweerdrfvaqpnkytlievqgeaesqeavvkvdsgpvrtvvkpcsld 180

Qy  181 PATQNLITNIFSKEMFKNAMTLMNLDVKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 patqnlitnifskemfknamtlmnldvkmplgkltkqqiargfealealeeamknptgd 240

Qy  241 GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 gqsleelsscfytviphnfgrsrpppinspdvlqakkdmllvladielaqtlqaapgeee 300

Qy  301 EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWKVNREG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ekveevphpldrdyqllrcqlqlldsgeseykaiqtylkqtgnsyrcpnlrhvwkvnreg 360

Qy  361 EGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGY 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 egdrfqahsklgnrrllwhgtnvavvaailtsglrimphsggrvgkgiyfasensksagy 420

Qy  421 VTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIEL 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 vttmhcgghqvgymflgevalgkehhitiddpslkspppgfdsviarggtepdpaqdiel 480

Qy  481 ELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEIHL 528
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 eldqgpvvvpqgppvqcpsfksssssfsqseyliykesqcrlrylleihl 528
```

RESULT 2
AAY51177
ID   AAY51177 standard; Protein; 533 AA.
XX
AC   AAY51177;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Murine PARP1 (long) homologue protein.
XX

KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP1.
XX
OS   Mus sp.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44290.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 67-69; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   murine PARP1 protein used in the method of the invention.
XX
SQ   Sequence 533 AA;

Query Match   99.6%; Score 2776.5; DB 21; Length 533;
Best Local Similarity   99.1%; Pred. No. 5.7e-238;
Matches 528; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```
Qy    1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 mapkrkasvqtegskkqrqgteeedsfrstaealraapadnrvirvdpscpfsrnpgiqv 60

Qy   61 HEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLEDAKK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 hedydctlnqtnignnnkfyiiqlleegsrffcwnrwgrvgevggqskmnhftcledakk 120

Qy  121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK-----VDSGPVRTVVK 175
         |||||||||||||||||||||||||||||||||||||||||||     ||||||||||||
Db  121 dfkkkfwektknkweerdrfvaqpnkytlievqgeaesqeavvkalspqvdsgpvrtvvk 180

Qy  176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKMPLGKLTKQQIARGFEALEALEEAMK 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 pcsldpatqnlitnifskemfknamtlmnldvkmplgkltkqqiargfealealeeamk 240

Qy  236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 nptgdgqsleelsscfytviphnfgrsrpppinspdvlqakkdmllvladielaqtlqaa 300

Qy  296 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWK 355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 pgeeeekveevphpldrdyqllrcqlqlldsgeseykaiqtylkqtgnsyrcpnlrhvwk 360
```

```
Qy   356 VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 vnregegdrfqahsklgnrrllwhgtnvavvaailtsglrimphsggrvgkgiyfasens 420

Qy   416 KSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPA 475
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ksagyvttmhcgghqvgymflgevalgkehhitiddpslkspppgfdsviargqtepdpa 480

Qy   476 QDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 qdieleldgqpvvvpqgppvvqcpsfksssfsqseyliykesqcrlrylleihl 533
```

RESULT 3
AAY51175
ID   AAY51175 standard; Protein; 533 AA.
XX
AC   AAY51175;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Human brain PARP3 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP3.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44288.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
PS   Claim 4; Page 57-59; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP3 protein used in the method of the invention.
XX
SQ   Sequence   533 AA;

Query Match   80.8%;   Score 2253.5; DB 21; Length 533;
Best Local Similarity 80.9%;   Pred. No. 1.8e-191;
Matches 432; Conservative 37; Mismatches 58; Indels 7; Gaps 4;
```

```
Qy     1 MAPKKASVQTEG--SKKQRQTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58
         ||||:||| ||||        | ||  ||| |||||||:|||||:||||||:||||||||
Db     1 mapkpkwvtegpekkkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt 60

Qy    59 QVHEDYDCTLNQTNIGNNNNKFYIIQLLEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117
         ||:|||||||||||   ||||||||||||||::  :|||:|||:||||||:|||||||
Db    61 qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqgskinhftrled 120

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTV---V 174
         |||:|||:||||| ||| ||||:||  || ||||:|||||:|  |||||| || |||  |
Db   121 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrgpvrtvtkrv 180

Qy   175 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQIARGFEALEALEEAM 234
         :||||||||:||||||||||||||  | | |||||||||||||||||||||||||:||:|
Db   181 qpcsldpatqkliitnifskemfkntnalmoldvkkmplgkiskqgiargfealealeeai 240

Qy   235 KNPTGDGQSLEELSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
         |:|  | |||||||:||||||||| ||:||||||| |||||:|||||:|||||||:||
Db   241 kgptdggqsleelshfytviphnfghsqpppinspeliqakkdmllvladielaqalqa 300

Qy   295 APGEEEKVEEVPHPLDRDYQLLRCQLOLQLLDSGESEYKAIQTYLKOTGNSYRCPNLRHVW 354
         : :  ||:|||||||||| |:|||:|||||||| ||||:|||| |:  ||:|:|||:|:|||
Db   301 v-seqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcptlqhiw 359

Qy   355 KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 414
         |||:|||||||||||||||| |||||||:||||||||||||||||||||||||||||||||
Db   360 kvnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiyfasen 419

Qy   415 SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
         |||||||| |:|| :|:|||||||||||:|||||:|:|||||||| |||||||| ||||||
Db   420 sksagyvtgmkcgahhvgymflgevalgrehhintdnpslksppppgfdsviarghtepdp 479

Qy   475 AQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
         | ||||||||| ||||| || ||||:|||||||:||||||:||||||||||:||:||
Db   480 tqdteleldgqqvvvpqgqvpcpefssstfsqseyliyqesqcrlryllevhl 533
```

RESULT 4
AAY51176
ID   AAY51176 standard; Protein; 540 AA.
XX
AC   AAY51176;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Human uterus type 2 PARP3 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP3.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44289.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -

XX
PS   Claim 4; Page 62-64; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD$^+$-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2HX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP3 protein used in the method of the invention.
XX
SQ   Sequence   540 AA;

Query Match   80.8%;   Score 2253.5; DB 21; Length 540;
Best Local Similarity 80.9%;   Pred. No. 1.9e-191;
Matches 432;   Conservative 37; Mismatches 58; Indels 7; Gaps 4;

Qy   1 MAPKRKASVQTEG--SKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58
         ||||:|||||||| ||||:|||||||:||||||||||:||: ||:|||||:|||||:||||
Db   8 mapkpkpwvqtegpekkkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt 67

Qy   59 QVHEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117
         ||:|| |||||||||:||||||||||||:: :||| ||||||||||||||:||||:|||
Db   68 qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftrled 127

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTV---V 174
         ||||| |||| ||||:||| || || ::||||||||| ::|:||||||||:||||| |
Db   128 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrgpvrtvtkrv 187

Qy   175 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234
         :|||||||| |||||||||||||| |:|| |||||||||||:|:||||||||||||||||
Db   188 qpcsldpatqklitnifskemfkntmalmdldvkkmplgkiskqqiargfealealeeal 247

Qy   235 KNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
         | |||||||||||||:||||:||||||:||:||||||:|||||||||||||||||||:||
Db   248 kgptdggqsleelsshfytvtiphnfghsqpppinspellqakkdmllvladielaqalqa 307

Qy   295 APGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVW 354
          :| |||||||||:||| |||||:||:|:|||||||:||:||||:::||::|||:|:|:|
Db   308 v-seqektveevphplrdrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcptlqhiw 366

Qy   355 KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 414
         |||:|||:|||||||||||:||||||||:||||||||||||||||||||||||||||||||
Db   367 kvnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiyfasen 426

Qy   415 SKSAGYVTTMHCGGHOVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
         ||||||| | ||| |:||||||||||| |||||| ||||||||||||| |||||||:|||||
Db   427 sksagyvigmkcgahhvgymflgevalgrehhintdnpslksppppfdsviarghtepdp 486

Qy   475 AQDIELELDGQPVVVPQGGPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEIHL 528
          :|:||||||| ||||||| | ||| ::| :|||||| :|:|||||||||:|:|:|
Db   487 tqdteleldgqqvvvpqggqpvpcpefssstfsqseyllyqesqcrlrylllevhl 540

RESULT 5
AAU29021
ID   AAU29021 standard; Protein; 533 AA.
XX
AC   AAU29021;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human PARP-3 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;

KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45590.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 105-107; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence   533 AA;

Query Match   80.4%;   Score 2241.5; DB 22; Length 533;
Best Local Similarity 80.5%;   Pred. No. 2.1e-190;
Matches 430;   Conservative 37; Mismatches 60; Indels 7; Gaps 4;

Qy   1 MAPKRKASVQTEG--SKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58
         ||||:|||||||| ||||:|||||||:||||||||||:||: ||:|||||:|||||:||||
Db   1 mapkpkpwvqtegpekkkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt 60

Qy   59 QVHEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117
         ||:|| |||||||||:||||||||||||:: :||| ||||||||||||||:||||:|||
Db   61 qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftrled 120

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTV---V 174
         ||||| |||| ||||:||| || || ::||||||||| ::|:||||||||:||:||| |
Db   121 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrapvrtvtkrv 180

Qy   175 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234
         :|||||||| |||||||||||||| |:|| |||||||||||:|:||||||||||||||||
Db   181 qpcsldpatqklitnifskemfkntmalmdldvkkmplgkiskqqiargfealealeeal 240

Qy   235 KNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQA 294
         | |||||||||||||:||||:||||||:||:||||||:|||||||||||||||||||:||
Db   241 kgptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladielaqalqa 300

Qy   295 APGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVW 354
          :| |||||||||:||| |||||:||:|:|||||||:||:||||:::||::|||:|:|:|

```
Db   301  v-seqektveevphpldrdyqlkcqlqlldsgapeykviqtyleqtgsnhrcptlqhiw 359
Qy   355  KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 414
          ||:||| |||||| ||||||||||||||||||||||||||||||||||||||||||
Db   360  kvnqegeedrfgahskignrkllwhgtnmavvaalltsglrimphsggrvgkgiyfasen 419
Qy   415  SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
          |||||| | |   | |||||||||||:||| ||||||||||||||||||| |||||
Db   420  sksagyvlgmkcgahhvgymflgevalgrehhintdnpslkspppgfdsviarghtepdp 479
Qy   475  AQDIELELDGQPVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
          :||||||||| |||  |::|||||||||||||||||||||||||||||:||||
Db   480  tqdteleldgqvvvpqgqpvpcpefssstfsqseyliyqesqcrlryllevhl 533


RESULT   6
AAY68834
ID   AAY68834 standard; Protein; 653 AA.
XX
AC   AAY68834;
XX
DT   16-MAY-2000 (first entry)
XX
DE   A poly(ADP-ribose) polymerase NAP protein of zea mays.
XX
KW   NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
KW   programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
KW   pest; drought; heat; fungi; nematode; seed-shatter.
XX
OS   zea mays.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..159
FT                   /note= "these residues are specifically claimed in
FT                   claim 18"
XX
PN   WO200004173-A1.
XX
PD   27-JAN-2000.
XX
PF   12-JUL-1999; 99WO-EP04940.
XX
PR   17-JUL-1998; 98US-0118276.
XX
PA   (PLB2 ) PLANT GENETIC SYSTEMS NV.
XX
PI   Babiychuk E, Kushnir S, De Block M;
XX
DR   WPI; 2000-182436/16.
DR   N-PSDB; AA260616.
XX
PT   Modulating cell death, growth and stress resistance in eukaryotes,
PT   specifically plants, used, e.g. to impart fungus or nematode resistance
PT   -
XX
PS   Claim 18; Page 92-95; 126pp; English.
XX
CC   The present sequence represents the NAP protein of Zea mays. This
CC   protein is a poly(ADP-ribose) polymerase (PARP) protein (also known as
CC   poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
CC   cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide
CC   sequences can be used for modulation of programmed cell death in
CC   eukaryotic cells. The method is used, specifically in plants, to induce,
CC   or protect against, programmed cell death, depending on the extent to
CC   which PARP activity is reduced. Reducing expression of endogenous NAP
CC   class PARP only is also used to modulate programmed cell death, to
CC   increase growth rate and to produce plant cells that are more tolerant
CC   of stress (cold, chemical treatments, pathogens, pests, drought, heat,
CC   etc., or during transformation). Particular applications are generation
CC   of plants that are resistant to fungi or nematodes; are male or female
CC   sterile; or have better seed-shatter properties. The methods are also
CC   used to improve growth of transformed plant cells (and derived calli or
CC   complete plants).
```

```
XX
SQ   Sequence   653 AA;

Query Match             25.8%;  Score 719.5;  DB 21;  Length 653;
Best Local Similarity   35.3%;  Pred. No. 6.7e-55;
Matches 190;  Conservative 91;  Mismatches 205;  Indels 53;  Gaps 17;

Qy    13  GSKKQRQGTEEDSFRS----TAEALRAAPADNRV--IRVDPSCPFSRNPGIQVHED-Y  64
          :|:| :|:|| : ||       ||| |: |||| :   ||| : ::| |      ||
Db   138  gakevikggdeevvkkekmvtatkkgaavldqhipdhikvnyhv------lqvgdeiy 190
Qy    65  DCTLNQTNIGNNNKFYIIQLLEE--GSRFFCWNRWGRVGEVGQSKMNHFTCLED-AKKD 121
          | |||||:|:||||||||||||  :  |||||| |||||||  :     ||||
Db   191  datlnqtnvgdnnnkfyiiqvlesdaggsfmvynrwgrvgrvgvrgqdklhgpsptrdqalye 250
Qy   122  FKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKVDSGPVRTVVKPCSLD 180
          :|| |||:|: | |::: |   ||||| :|
Db   251  fegkfhnktnnhwsdrknfkcyakkytwiemdygeteke-----iekssitdqiketkle 305
Qy   181  PATQNLJTNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 240
          :|| |||||||| :  :| |  ::| :   ||||| :    |
Db   306  triaqfislicnismmkqrmveignyaeklplgklrkatilkgyhvlkrisdvisk--ad 363
Qy   241  GQSLEELSSCFYTVIPHNFG--RSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGE 298
          :|: |||||||||:|| |:||  :| :| ::||
Db   364  rrhleqitgefytviphdfgfrkmrefiidtpqklkaklemvealgeieiatkll----- 418
Qy   299  EEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT--GNSYRCPNLRHVWK 355
          :   :| ||:|| :  :||| ::||
Db   419  -eddssdqddplyarykqlhcdftpleadsdeysmiksylrnthgkthsgytvdivqifk 477
Qy   356  VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFA 411
          |:|:||| |||| ||:|| : ||||| : ||:||||||||||
Db   478  vsrhgeterfgkfastrnrmllwhgsrlsnwagilsgglrlappeapvtgymfgkgvyfa 537
Qy   412  SENSKSAGYVTTMHCGGHQV---GYMFIGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 468
          |||| :       : ||
Db   538  dmfsksany-----cyaseacrsgvllcevalgdmnellnadydannlpkgklrskgvg 592
Qy   469  QTEPDPAQDIELELDGQPVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIH 527
          : ||  | :: :|: :| :| ::|: ::|| : ||:||||||| ::| :
Db   593  qtapnm---veskvaddgvvvplgepkqeps-krggllyneylvynvdqirmyvlhvn 647


RESULT   7
AAR99642
ID   AAR99642 standard; Protein; 1013 AA.
XX
AC   AAR99642;
XX
DT   31-OCT-1996 (first entry)
XX
DE   Poly(ADP-ribose) polymerase contg. DNA-binding domain.
XX
KW   PARP; DNA-binding domain; poly(ADP-ribose) polymerase; gene therapy;
KW   tumour treatment; DNA repair; over-expression.
XX
OS   Homo sapiens.
XX
PN   WO9618737-A2.
XX
PD   20-JUN-1996.
XX
PF   15-DEC-1995;   95WO-DE01817.
XX
PR   16-DEC-1994;   94DE-4444949.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI   Buerkle A, Kuepper J, Zur Hausen H;
XX
DR   WPI; 1996-300654/30.
```

```
DR   N-PSDB: AAT13732.
XX
PT   Vectors contg. insert encoding DNA-binding domain of
PT   poly(ADP-ribose) polymerase  -  useful for gene therapy, esp. of
PT   tumours
XX
PS   Disclosure; Fig 1; 22pp; German.
XX
CC   The present sequence is that of a poly(ADP-ribose) polymerase (PARP)
CC   contg. a DNA-binding domain (DBD).  It is encoded by AAT13732, from
CC   which a fragment comprising nucleotides 67-1220 (-29 to +1127) encoding
CC   the DBD can be inserted into vectors which are used for gene therapy.
CC   Over-expression of the DBD inhibits the DNA repair function of PARP, so
CC   the vectors are useful for gene therapy or tumours, esp. in combination
CC   with conventional chemo- and/or radiotherapy.
XX
SQ   Sequence 1013 AA;

Query Match            25.4%;     Score 707.5;  DB 17;  Length 1013;
Best Local Similarity  34.0%;     Pred. No. 1.5e-53;
Matches 183;  Conservative 101;  Mismatches 216;  Indels 39;  Gaps 17;

Qy    1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
         ::||: ||: :|  |  : |||||||||||||||||||||||||||||||||||||||||
Db  492 vaprgksga--alskkaskgqvkeeginksekrmkltikggaavdpdsglehsahvlekg 548

Qy   61 HEDYDCTLNQTNIGNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
         ::: || :| :| :|||||||||||| |:  :|:|||||||:| |:||||||:| :|
Db  549 gkvfsatiglvdivkgtnsyykiqlleddkenrywifrswgrvgtvigsnkleqmpsked 608

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQRAVVK--VDSGPVRTVVK 175
         |:: |: :|:| :| :| :| | | :||:|: :| | :|  : :|  | | :|||::|
Db  609 aiehfmklyeektgnawhskn-ftkypkkfiypleid-yqqdeeavkkltvngtkskipk 666

Qy  176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 235
         |       :| :: |: ::|:| || :|:|| ::||:||||:|||:||| |  :  : |
Db  667 p------vgdllkmifdvesmkkamveyeidlqkmplgklskrqiqaaysllsevqqavs 720

Qy  236 NPTGDGQGSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
         :   : | |: :: :||| |||:||| :||| ||:|| ||| || ||| || : :|:
Db  721 qgssdsqil-dlsnrfytliphdfgmkkppllnnadsvqakvemldnlldievaysllrg 779

Qy  296 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLR 351
         :| ::  : :  :  : ::|::: : :| : :::| : ::|: :    | :: : |  :
Db  780 gsddssk------dpidvnyeklkdikvvdrdseeaeiirkyvknthatthnaydl-evi 833

Qy  352 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKG 407
         :  | ||| | |: |: | |:||||||||  :  | |  ::| || ||     || :|||
Db  834 difkieregecqrykpfkqlhnrrllwhgsrtnfagilsqgiriappeapvtgymfgkg 893

Qy  408 IYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIAR 467
         |||| : ||:| : | |   : :|  :::|||||| : || | |     ::|:   | |
Db  894 iyfadmvsksanychtsq-gdpiglilligevalgnmyelk-hashisklpkgkhsvkgl 950

Qy  468 GQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEI 526
         | | |||:  : |:| |  || | :  :  | : :  |:  : ||::| ::: :|:||| :
Db  951 gkttpdpsani--sldg--vdvplgtgis-sgvndtsllyneyivydiaqvnlkyllki 1004
```

RESULT 8

```
AAU29022
ID   AAU29022 standard; Protein; 522 AA.
XX
AC   AAU29022;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Mouse PARP-2 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Mus musculus.
XX
PN   WO20164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I,  Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45597.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 109-111; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP.  PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis.  The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues.  They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia) , inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes.  The present sequence is a PARP protein,
CC   the cdNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence  522 AA;

Query Match            25.3%;     Score 707;  DB 22;  Length 522;
Best Local Similarity  35.4%;     Pred. No. 6.2e-54;
Matches 190;  Conservative 88;  Mismatches 195;  Indels 64;  Gaps 21;

Qy   20 GTEEEDSFRSTAEALRA------APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTLNQ 70
         | : || :::|:  : |:      || ::| : | :  :      :: :||:: | :|:|
Db   12 ggkdadrtkdnrdsvktllikgkap------vdpecaaklgkahvycegddvydvmlnq 64

Qy   71 TNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGQSKMNHFTC--LEDAKKDFKKK 125
         || |:||||| |||||| |:  | | :|||||| : :|:: |    : || : |:|
Db   65 tnlqfnnnkyyliqlleddagrnfsvwmrwgrvgktgqhsl--vtcsgdlnkakelfqk 122

Qy  126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQ-EAVVKVDSGPVRTVVKPCS-LDPAT 183
         |::|| |:||:|: |:  | |   : : :|| : :  : : |   : :: | :  || :|
Db  123 fldktknnwedrenfekvpgkydmiqmdyaastqdesktkeee----tlkpesqldlrv 177

Qy  184 QNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD-GQ 242
         |:|: ::   ||:: |  :: : :  : :||||  :|  |::| :|  : :: :    |:
Db  178 qellklicnvqtmeemniemkydtkrapigkltvaqikagyqslkkiedcir--agqhgr 235

Qy  243 SLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEK 302
         :: |   :::| ||::|| | ||::  :|: :| :  : |:|: ::|:  ||      :
Db  236 alveacnefytriphdfglsippvirtekelsdkvkllealgdieialklvks----erqg 292

Qy  303 VEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------NLRHVWK 355
         :| :|||:: ::  |:   : ||:  : :: : : :: : :      :   | : :| :
Db  293 le--hpldqhyrnlhcairpldhesneffkvisqylgst---hapthkdytmtlldvfe 345
```

```
Qy      1 MAPRKRASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          :||:|: ::   :| :: :|::  |:: :      ::: :::::|   ||: |:::|:::
Db    493 vaprgksga--alskkskgqvkeeginksekrmkltlkggaavdpdsglehsahvlekg 549

Qy     61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          |||||| :|::  |||| | |:: : ::     |:: |:::|::|: ::|:|
Db    550 gkvfsatiglvdivkgtnsyyklqlleddkenrywifrswgrvgtcvigsnkleqmpsked 609

Qy    118 AKKDFKKKFWEKITKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVK 175
          |  || |:|::     :|   | :  |:: :| |: |: :: ::|  | | ||: |||| 
Db    610 aieqfmklyeektgnawhskn-ftkypkkfypleid-yggdeeavkkltvnpgtkskipk 667

Qy    176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 235
          || ::|| ||   ::||||||:|:|| :: :::|::|:: | ||:|:|:| ::|::|
Db    668 p-----vqdlikmifdvesmkkamveyeidlqkmplgklskrqigaaysilsevqqavs 721

Qy    236 NPTGDGQSLEELSSCFYIVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
          :: : | :|| |  | |:|  |:|  |:|| ::: : :|:|:| :||::|:|::| :: 
Db    722 qgsdsqil-disnrfytliphdfgmkkpplnnadsvqakvemldnlldievaysllrg 780

Qy    296 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--H 352
          : : | |    | |:| ::|:|::|:  : :|| :|    :|::|  ||:| :: ::| H
Db    781 gsddssk----qpidvnyeklktdikvvdrdseeaeilrkyvknthatthsaydlevid 835

Qy    353 VWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGI 408
          ::|:|||| |||  :|||| |||||||||| ::|||:| |::|||  :: ||||||||
Db    836 ifkieregecqrykpfkqlhnrrllwhgsrttnfagilsqglriappeapvtgymfgkgi 895

Qy    409 YFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 468
          ||| ||||||| ||| |  ||:|||||||||  || | ::| |||:| | :: ::| |
Db    896 yfadmvsksanyyhtsq--gdpiglilgevalgnmyeik-hashisrlpkgkhsvkglg 952

Qy    469 QTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLLYKESQCRLRYLLEI 526
          ||| : |  ::| |: | || :: |: : |   |::: :|:: |:|: ::| :|||||
Db    953 kttpdpsani--sldg--vdvplgtgis-sgvidtsllyneyivydiagvnlkyllki 1005
```

```
Qy    356 VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFA 411
          : |||| | :|||||| |:  |||| ||  |:|: |:| :::| :: ||||||||||||
Db    346 vekegekeafr--edlpnrmllwhgsrlsnwvgilshglrvappeapitgymfgkglyfa 403

Qy    412 SENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTE 471
          | ||| |    | |  | ::| || :  | |||||:|::::|:  |:|  |:: ||: :
Db    404 dmssksanycfasrl--kntgllllsevalgqcnelleanpkaqgllrgkhstkgmqkma 461

Qy    472 PDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSF-SQSEYLLYKESQCRLRYLLEI 526
          | |  |:::|  :  |:|  :       ||: ||| | |:|:|:||:||:|||||||:|
Db    462 pspahfi--tings--tvplgpasdtgilnpegytlnynefivyspqvrmryllki 514
```

RESULT 9
AAY58043
ID   AAY58043 standard; peptide; 1014 AA.
XX
AC   AAY58043;
XX
DT   14-MAR-2000 (first entry)
XX
DE   Human poly (ADP-ribose) polymerase protein sequence.
XX
KW   Poly (ADP-ribose) polymerase; DNA excision repair; DNA damage; PARP;
KW   carcinogen; cancer therapy; base modification; human; cancer; UvrA; UvrB.
XX
OS   Homo sapiens.
XX
PN   WO9958971-A2.
XX
PD   18-NOV-1999.
XX
PF   13-MAY-1999;　99WO-CA00444.
PR   13-MAY-1998;　98US-0078347.
XX
PA   (LEXX/) LE X.
PA   (WEIN/) WEINFELD M.
PA   (XING/) XING J Z.
XX
PI   Le X, Weinfeld M, Xing JZ;
XX
DR   WPI; 2000-062317/05.
XX
PT   Quantifying nucleic acid modifications useful for identifying and
PT   detecting exposure to carcinogens, in early risk assessment for cancer,
PT   and in monitoring cancer therapy -
XX
PS   Example 6; Fig 9; 65pp; English.
XX
CC   This is the human poly (ADP-ribose) polymerase protein (PARP) sequence.
CC   PARP is a 116kD DNA-binding protein that binds tightly to single-strand
CC   and double-strand breaks. The PARP protein and Escherichia coli UvrA and
CC   UvrB proteins (see AAY58041-Y58042) are used in a method for the
CC   detection and quantitation of DNA base damage. The method involves
CC   quantifying at least one modification of interest in a nucleic acid
CC   sequence through the used of capillary electrophoresis, laser induced
CC   fluorescence and a fluorescently labeled polypeptide (e.g. fluorescently
CC   labeled PARP or fluorescently labeled anti-PARP antibody) that is
CC   specific for the modification. The new method is useful for quantifying
CC   nucleic acid modifications, especially a mutation (including deletions,
CC   insertions and substitutions), mismatch, DNA adduct, or strand break
CC   (single or double). The method is useful for identifying and detecting
CC   exposure to carcinogens, and in early risk assessment for cancer, and is
CC   also useful in monitoring cancer therapy.
XX
SQ   Sequence   1014 AA;

Query Match   25.2%; Score 702.5; DB 21; Length 1014;
Best Local Similarity 33.8%; Pred. No. 4.3e-53;
Matches 182; Conservative 102; Mismatches 217; Indels 37; Gaps 17;

RESULT 10
AAU29019
ID   AAU29019 standard; Protein; 1014 AA.
XX
AC   AAU29019;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human PARP-1 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45582.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes

PT   inhibits human PARP -
XX
PS   Example 13; Page 97-101; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia) , inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence 1014 AA;
XX

Query Match   25.2%; Score 702.5; DB 22; Length 1014;
Best Local Similarity   33.8%;   Pred. No. 4.3e-53;
Matches 182; Conservative 102; Mismatches 217; Indels 37; Gaps 17;

Qy   1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          ::|: |: | : | :| | : | ||::| | : | | | -| | :|
Db   493 vaprgksga--alskkskggvkeeginksekrmkltkggaavdpdsglehsahvlekg 549

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          :: || | : : | :|||:||||||| :|||||||||||| || :| :
Db   550 gkvfsatlglvdivkgtnsyyklqlleddkenrywifrswgrvgtvigsnkleqmpsked 609

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVK 175
          | | |:: || || || | | | || | : || | : |: : | | : |::
Db   610 aieqfmklyeektgnawhskn-ftkypkkfypleid-yggdeeavkkltvnpgtkskipk 667

Qy   176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 235
          | :||| :|| || :||| ||:| |||||| |:||: | ::| :::|
Db   668 p------vqdlikmifdvesmkkanveyeidlqkmplgkiskrqiqaaysilsevqqavs 721

Qy   236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
          : | | |: | | |||| ||||:|| ||| || ||:|:|||::| :|| ::| :: ||
Db   722 qgssdsqil-dlsnrfytliphdfgmkkpplinnadsvqakvemldnlldievaysllrg 780

Qy   296 PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--H 352
          |:: | :| | :| :|:|| |:| | || | | : :| ||:| || :| :|
Db   781 gsddssk-----dpidvnyeklktdikvvdrdseeaeilrkyvknthatthsaydlevid 835

Qy   353 VWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGI 408
          ::|: ||:|| |: || ||||||||||: ::|| ||:|| |: | :||||:||
Db   836 ifkieregecgrykpfkqlhnrrllwhgsrttnfagilsqglriappeapvtgymfgkgi 895

Qy   409 YFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 468
          ||| : ||| : | | | || ::||||||||| : || || ||:| |:|::|:|
Db   896 yfadmvsksanyyhtsq--gdpiglillgevalgnmyelk-hashisrlpkgkhsvkglg 952

Qy   469 QTEPDPAQDIELELDGQPVVPQGGPPVQCPSFKSSSSFSQSEYLIIYKESQCRLRYLLEI 526
          ::||||: | | : | | | | :: |: | || ::|: ||:|:|||:|
Db   953 kttpdpsani--sldg--vdvplgtgis-sgvidtsllyneyivydiaqvnlkylki 1005

RESULT 11
AAB66296
ID   AAB66296 standard; Protein; 1014 AA.
XX
AC   AAB66296;
XX
DT   05-APR-2001 (first entry)
XX
DE   Human tankyrase2 related protein sequence SEQ ID NO: 137.
XX

KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US17827.
XX
PR   29-JUN-1999; 99US-0141582.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63954.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders -
XX
PS   Example 2; Page 207-210; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC   polyADP-ribosylation activity and is involved in the modification of
CC   TRF1, which is a telomere-specific binding protein. The regulation of
CC   telomere length, in which TRF1 has a role, is linked to ageing and
CC   cancer. The sequences are useful in the treatment of cancers and
CC   inflammatory disorders.
XX
SQ   Sequence   1014 AA;
XX

Query Match   25.2%; Score 702.5; DB 22; Length 1014;
Best Local Similarity   33.8%;   Pred. No. 4.3e-53;
Matches 182; Conservative 102; Mismatches 217; Indels 37; Gaps 17;

Qy   1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          ::|: |: | : | :| | : | ||::| | : | | | -| | :|
Db   493 vaprgksga--alskkskggvkeeginksekrmkltkggaavdpdsglehsahvlekg 549

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          :: || | : : | :|||:||||||| :|||||||||||| || :| :
Db   550 gkvfsatlglvdivkgtnsyyklqlleddkenrywifrswgrvgtvigsnkleqmpsked 609

Qy   118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVK 175
          | | |:: || || || | | | || | : || | : |: : | | : |::
Db   610 aieqfmklyeektgnawhskn-ftkypkkfypleid-yggdeeavkkltvnpgtkskipk 667

Qy   176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 235
          | :||| :|| || :||| ||:| |||||| |:||: | ::| :::|
Db   668 p------vqdlikmifdvesmkkanveyeidlqkmplgkiskrqiqaaysilsevqqavs 721

Qy   236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
          : | | |: | | |||| ||||:|| ||| || ||:|:|||::| :|| ::| :: ||
Db   722 qgssdsqil-dlsnrfytliphdfgmkkpplinnadsvqakvemldnlldievaysllrg 780

Qy   296 PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--H 352
          |:: | :| | :| :|:|| |:| | || | | : :| ||:| || :| :|
Db   781 gsddssk-----dpidvnyeklktdikvvdrdseeaeilrkyvknthatthsaydlevid 835

Qy   353 VWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGI 408
          ::|: ||:|| |: || ||||||||||: ::|| ||:|| |: | :||||:||
Db   836 ifkieregecgrykpfkqlhnrrllwhgsrttnfagilsqglriappeapvtgymfgkgi 895

Qy   409 YFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 468
          ||| : ||| : | | | || ::||||||||| : || || ||:| |:|::|:|
Db   896 yfadmvsksanyyhtsq--gdpiglillgevalgnmyelk-hashisrlpkgkhsvkglg 952

```
Qy  469  QTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSFSQSEYLIYKESQCRLRYLLEI 526
         :|:|||::| ||| |.| | :|  :| :||||:||:  || |:|||:||::
Db  953  ktttpdpsani--sldg--vdvplgtgis-sgvidtsllyneyivydiaqvnlkylikl 1005

RESULT 12
AAY33699
ID  AAY33699 standard; Protein; 1014 AA.
XX
AC  AAY33699;
XX
DT  18-JAN-2000   (first entry)
XX
DE  Human poly(ADP-ribose)polymerase protein.
XX
KW  Gene therapy; poly(adenosine diphosphate-ribose) polymerase; treatment;
KW  PARP; antitumor; nuclear DNA repair; proliferating cell; DNA damage;
KW  protection; genomic instability; cancer; prevention; human.
XX
OS  Homo sapiens.
XX
XX  DE19808889-A1.
PN
XX  09-SEP-1999.
PD
XX
PF  03-MAR-1998; 98DE-1008889.
XX
PR  03-MAR-1998; 98DE-1008889.
XX
PA  (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI  Buerkle A, Meyer R;
XX
DR  WPI; 1999-509563/43.
DR  N-PSDB; AAZ23799.
XX
PT  New gene therapy vector expressing poly(adenosine
PT  diphosphate-ribose) polymerase for treating or preventing tumors -
XX
PS  Disclosure; Fig 2; 12pp; German.
XX
CC  This invention describes a novel vector (A), suitable for gene therapy,
CC  which contains a DNA insert (I) that encodes a practically complete
CC  poly(adenosine diphosphate-ribose) polymerase (PARP) which has antitumor
CC  activity. PARP is a nuclear DNA repair enzyme (activated by strand
CC  breakage), that allows recovery of proliferating cells from the toxic
CC  effects of DNA damage (caused by alkylating or oxidizing agents or
CC  radiation), and protects cells against such damage or other causes
CC  of genomic instability. (A) are used for treatment of cancer and for
CC  cancer prevention in subjects at high risk (e.g. those with
CC  tumor-associated genetic defects). Tumor cells treated with (A) show an
CC  increased tendency to die when treated with radiation or chemotherapeutic
CC  agent, and in cells that survive this treatment they inhibit genomic
CC  instability, so should reduce the likelihood of further development of
CC  both malignant cells and resistance to chemotherapy. This sequence
CC  represents the human poly(ADP-ribose) polymerase described in the
CC  invention.
XX
SQ  Sequence  1014 AA;

Query Match      25.2%;  Score 701.5;  DB 20;  Length 1014;
Best Local Similarity  34.0%;  Pred. No. 5.3e-53;
Matches  183;  Conservative  99;  Mismatches 218;  Indels 39;  Gaps  17;

Qy    1  MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
         :||: || :| :| :| || :|| ::  :::|   | :  | :
Db  493  vaprgksga--alskkskgqvkeeginksekrmkltlkggaavdpdsglehsahvlekg 549

Qy   61  HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
         :||||  |::: |||||| || ||  | ::|||| || | |:|
Db  550  gkvfsatlglvdivkgtnsyyklqllleddkenrywifrswgrvgtvigsnkleqmpsked 609
```

```
Qy  118  AKKDFKKFWEKTKNKWEERDRFVAQPNKYTLEVQGEAESQEAVVK--VDSGPVRTVVK 175
         | :||| ::| ||| | :||| || :| | :||:||||: ::|  ::|:: :|
Db  610  aiehfmklyeektgnawhskn-ftkypkkfypleid-yggdeeavkkltvnpgtkskipk 667

Qy  176  PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQOIARGPEALEALEEAMK 235
         |  :||  || ||| |||    || ||  :||::|||||:|||::
Db  668  p-----vqdlikmifdvesmkkamveyeidlqkmplgkslskrqiqaaysilsevqqavs 721

Qy  236  NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
         :|| | | |||:|| ||| ||  ||| :|| :: |:| | :: || | ||||:
Db  722  qgssdsqil-dlsnrfytliphdfgmkpplinnadsvqakvemlnlldievaysllrg 780

Qy  296  PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYRCPNLR 351
         ::| |:  :  :|: ||||   |:||| |  || | :|:|:
Db  781  gsddssk-----dpidvnyeklktdikvvdrdseeeaeilrkyvknthatthaaydl-evi 834

Qy  352  HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKG 407
         :|:|||| |:: | :|| ||:| ::|||| | | ||| :  :| :| |||
Db  835  difkieregecqrykpfkqlhnrrllwhgsrtnfagilsqglriappeapvtgymfgkg 894

Qy  408  IYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEEHITIDDPSLKSPPPGFDSVIAR 467
         |||| |||| : ||  :|: ::|| | : ||||||:
Db  895  lyfadmvsksanychtsq--gdpiglilgevalgnmyelk-hashislpkgkhsvkgl 951

Qy  468  GQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI 526
         :|:|||: | | || :| : || |:||| :|| :| | ::|||| ::  || |:
Db  952  gkttpdpsani--sldg--vdvplgtgts-sgvndtsllyneytvydiaqvnlkylikl 1005

RESULT 13
AAB60693
ID  AAB60693 standard; Protein; 521 AA.
XX
AC  AAB60693;
XX
DT  11-MAY-2001   (first entry)
XX
DE  Human poly(ADP-ribose) synthetase sbhPARS2.
XX
KW  Human; poly(ADP-ribose) synthetase; sbhPARS2; vaccine; drug screening;
KW  ischaemic disorder; cerebral ischaemia; cardiac ischaemia;
KW  myocardial infarction; stroke; inflammation; autoimmune disease;
KW  diabetes; multiple sclerosis; neurodegenerative disease;
KW  Parkinson's disease; Alzheimer's disease; chromosome localisation.
XX
OS  Homo sapiens.
XX
PN  WO200112645-A1.
XX
PD  22-FEB-2001.
XX
PF  10-AUG-2000; 2000WO-US21775.
XX
PR  12-AUG-1999; 99US-0373441.
XX
PA  (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI  Barone F, Field J, Kabnick KS, Li X, McQueney MS, Zhu Y;
XX
DR  WPI; 2001-211196/21.
DR  N-PSDB; AAF59996.
XX
PT  New human polypeptide of the polyADPribose synthetase family for use in
PT  diagnostic assays and for screening modulators used for preventing and
PT  treating inflammation, autoimmune disease and Alzheimers disease -
XX
PS  Claim 1; Page 21; 30pp; English.
XX
CC  This sequence represents a novel human poly(ADP-ribose)
CC  synthetase, sbhPARS2. The invention also relates to fragments,
CC  variants and sequences with at least 95% identity to the sbhPARS2
CC  protein or nucleotide sequence; expression systems and host cells
```

CC   comprising an sbhPARS2 nucleic acid sequence; the recombinant expression
CC   of sbhPARS2; and an antibody specific for sbhPARS2. sbhPARS2 proteins
CC   and nucleotides are useful as vaccines for inducing an immunological
CC   response in a mammal. The sbhPARS2 protein is useful for identifying
CC   compounds which inhibit or stimulate its activity or expression level.
CC   Such agonists and antagonists of sbhPARS2 are useful for treating human
CC   diseases including ischaemia and ischaenic tissue injury (e.g., cerebral
CC   and cardiac ischaemia, myocardial infarction, stroke), inflammation,
CC   autoimmune diabetes (e.g. diabetes, multiple sclerosis) and
CC   neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's
CC   disease). sbhPARS2 nucleic acids are useful as diagnostic reagents for
CC   detecting mutations in the associated gene; as hybridisation probes
CC   to isolate full-length sbhPARS2 cDNAs and sbhPARS2 genomic clones; and
CC   for chromosome localisation studies. The sbhPARS2 protein is also useful
CC   as an immunogen to produce antibodies for therapeutic use. sbhPARS2
CC   proteins, nucleotides and antibodies are also useful in screening methods
CC   for detecting the effect of added compounds on the production of mRNA and
CC   protein in cells.
XX
SQ   Sequence   521 AA;

Query Match   25.1%;   Score 700.5;   DB 22; Length 521;
Best Local Similarity   34.6%;   Pred. No. 2.3e-53;
Matches   185; Conservative   92; Mismatches   212; Indels   45; Gaps   17;

Qy   13 GSKKQRQGTEEEDSFRSTAEALRA-APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTL 68
           | | ||::: :| || :: || ||| || :|| ::||
Db   5 ggkankdrtedkqdesvkallikgkap------vdpectakvgkahvycegndvydvml 57

Qy   69 NQTNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFK 123
           |||| |||:::|||:|:|| :: :| |||||||||||||:||::| | |: ::|
Db   58 nqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhsl--vacsgnlnkakeifq 115

Qy   124 KKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCS-LDPA 182
           |||| ||||| |:| | | : | :|||:| :: : |:   ::| ||: ||: | |
Db   116 kkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpesqldlr 175

Qy   183 TQNLTNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD-G 241
           | | ||| :| | | | | | || |:||:| : : ||:: :|: : :| :
Db   176 vgeliklicnvqameemmmemkynttkkaplgkltvaqikagyqslkkiedcir--agqhg 233

Qy   242 QSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEE 301
           :: || | | | || | : : :| | |||:|:|:| | | ::| | |
Db   234 ralmeacnefytriphdfgirtpplirtqkelsekiqllealgdieiaikl------vkt 287

Qy   302 KVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT--GNSYRCPNLRHVWKVNR 358
           |:: |||| || :||| :| | :| : | | | ::   :||| | |: ::|
Db   288 elqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlifevek 347

Qy   359 EGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASEN 414
           || | | || : | || | ||: :|: : | |:|| :| |||:| |||:| ::
Db   348 dgekeafr--edlhnrmllwhgsrmsnwwgilshglriappeapitgymfgkgiyfadms 405

Qy   415 SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
           | || | |||| |: : | |:|||: ::| || ::| |:| | | | |
Db   406 sksanycfasrl--kntglllsevalgqcnelleanpkaegllqgkhstkglgkmapss 463

Qy   475 AQDIELELDGQPVVPQGPPVQCPSFKSSSF--SQSEYLIYKESQCRLRYLLEI 526
           : |||| | || | | ||:|||:| | | : :|:| :||:|| |:|
Db   464 ahfv--tlngs--tvplgpasdtgilnpdgytlnyneyivynpqvrmrryllkv 513

RESULT 14
AAY51174
ID   AAY51174 standard; Protein; 570 AA.
XX
AC   AAY51174;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Human brain PARP2 protein.
XX

KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB: AAZ44287.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 52-54; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP2 protein used in the method of the invention.
XX
SQ   Sequence   570 AA;

Query Match   25.1%;   Score 700.5;   DB 21; Length 570;
Best Local Similarity   34.6%;   Pred. No. 2.7e-53;
Matches   185; Conservative   92; Mismatches   212; Indels   45; Gaps   17;

Qy   13 GSKKQRQGTEEEDSFRSTAEALRA-APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTL 68
           | | ||::: :| || :: || ||| || :|| ::||
Db   54 ggkankdrtedkqdesvkallikgkap------vdpectakvgkahvycegndvydvml 106

Qy   69 NQTNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFK 123
           |||| |||:::|||:|:|| :: :| |||||||||||||:||::| | |: ::|
Db   107 nqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhsl--vacsgnlnkakeifq 164

Qy   124 KKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCS-LDPA 182
           |||| ||||| |:| | | : | :|||:| :: : |:   ::| ||: ||: | |
Db   165 kkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpesqldlr 224

Qy   183 TQNLTNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD-G 241
           | | ||| :| | | | | | || |:||:| : : ||:: :|: : :| :
Db   225 vgeliklicnvqameemmmemkynttkkaplgkltvaqikagyqslkkiedcir--agqhg 282

Qy   242 QSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEE 301
           :: || | | | || | : : :| | |||:|:|:| | | ::| | |
Db   283 ralmeacnefytriphdfgirtpplirtqkelsekiqllealgdieiaikl------vkt 336

Qy   302 KVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRCPNLRHVWKVNR 358
           |:: |||| || :||| :| | :| : | | | ::   :||| | |: ::|
Db   337 elqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlifevek 396

CC after heart valve replacement, aneurysm resectioning and heart
CC transplantation), revascularization of critically constricted coronary
CC arteries (e.g. after PCTA or by-pass operations) or peripheral arteries
CC (e.g. leg arteries), treating acute myocardial ischemia and damage during
CC or after its mechanical or drug-induced lysis and treating tumors and
CC their metastasis, sepsis and septic shock, inflammatory and rheumatic
CC disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)
CC inhibit PARP (i.e. the known form designated PARP), they especially
CC selectively and strongly inhibit PARP homologs, specifically the homolog
CC PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.
CC with K_i values of 1-20 nM) and high selectivity for PARP2 relative to
CC PARP1 (generally by a factor of more than 5).
XX
SQ Sequence 570 AA;

```
Query Match            25.1%;  Score 700.5;  DB 22;  Length 570;
Best Local Similarity  34.6%;  Pred. No. 2.7e-53;
Matches  185;  Conservative  92;  Mismatches  212;  Indels  45;  Gaps  17;

Qy    13 GSKKQRQGTEEDSFRSTAEALRA-APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTL  68
            ||:   |:    :||:::   |  :             :|| ||   :  : |||   |
Db    54 ggkankdrtedkqdesvkallkgkap------vdpectakvgkahvycegndvydvml 106

Qy    69 NQTNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFK 123
            |||| ||:||:||:||||::||   | | | ||||| | ||| ||    |:    ||:
Db   107 nqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhsl--vacsqnlnkakeifq 164

Qy   124 KKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCS-LDPA 182
            |||  |:||:||  |:| ::|  |  :  :   ::    ||    |:::  :|   |: :
Db   165 kkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslkspxpesqldir 224

Qy   183 TQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD-G 241
            .||::  :   ||:::   :: :  :||:|:|||| :| |:|: :  | |:    |  | |
Db   225 vqeliklicnvqameemmmemkyntkkaplgkltvaqikagyqslkkiedcir--agqhg 282

Qy   242 QSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEE 301
            ::  : ::|||:| ||:|| | |  |   :  ::|:: |::| :|:|| :|:        .
Db   283 ralmeacnefytriphdfglrtpplirtqkelsekiqlleaigdieiaikl------vkt 336

Qy   302 KVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRCPNLRHYWKVNR 358
            :| :  :| |:   : |  |:| ||:||:||: || ||:|   ||:|||:|:||||:
Db   337 elqspehpldqhyrnlhcalrlpldhesyefkvisqylqsthapthsdytmtlldifevek 396

Qy   359 EGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFASEN 414
            :|:: :|:   ||||:|||||| :    |  |:|||||:||    ||||:|||||||:
Db   397 dgekeafr--edlhnrmllwhgsrmsnwvgilshglriappeapitgymfgkglyfadms 454

Qy   415 SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
            ||||  | | :   :|||:  : ||| | | :||||||||| :   | |:|||:  |:||
Db   455 sksanycfasrl--kntglllsevalqgcnelleanpkaeglloqkhstkglgkmapss 512

Qy   475 AQDIELELDGQPVVPQGPPVQCPSFKSSSF--SQSEYLIYKESQCRLRYLLEI 526
            | | :|  :|:     |      ||||||   ||::|:  ||  :  |||||:
Db   513 ahfv--tlngs--tvplgpasdtgilnpdgytlnyneyivynpnqvrmryllkv 562
```

Search completed: August 29, 2002, 07:57:21
Job time: 298 sec

RESULT 15
AAB11480
ID   AAB11480 standard; Protein; 570 AA.
XX
AC   AAB11480;
XX
DT   02-MAR-2001 (first entry)
XX
DE   Human brain poly-ADP-ribose-polymerase protein.
XX
KW   Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
KW   cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
KW   anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
KW   antiinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
KW   2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
KW   neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
KW   Parkinson's disease; ischemic damage; microinfarction; sepsis;
KW   diabetes mellitus.
XX
OS   Homo sapiens.
XX
PN   DE19921567-A1.
XX
PD   16-NOV-2000.
XX
PF   11-MAY-1999; 99DE-1021567.
XX
PR   11-MAY-1999; 99DE-1021567.
XX
PA   (BADI) BASF AG.
XX
PI   Lubisch W, Sadowski J, Kock M, Hoeger T;
XX
PN   WPI; 2001-032983/05.
XX
DR   N-PSDB; AAC82090.
XX
PT   Drugs for inhibiting PARP or especially homologous enzymes comprising
PT   4-substituted phthalazinone derivatives, useful e.g. for treating
PT   neurodegenerative disease, ischemic damage, tumors or diabetes -
XX
PS   Example A; Page 12-13; 14pp; German.
XX
CC   This invention describes novel 4-substituted 2H-phthalazin-1-one
CC   derivatives (I) which are used for the treatment or prophylaxis of
CC   diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;
CC   EC 2.4.2.30) activity. The products of the invention have nootropic,
CC   neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,
CC   cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,
CC   immunosuppressive, antiinflammatory, antirheumatic, antiarthritic,
CC   antidiabetic. (I) are especially used for treating or preventing
CC   neurodegenerative disease or neuronal damage (specifically associated
CC   with ischemia, trauma or massive bleeding, especially apoplexy or
CC   spinal-cranial trauma; or Alzheimer's disease, Huntington's disease or
CC   Parkinson's disease), treating or preventing ischemic damage
CC   (specifically renal damage after renal ischemia or during and after
CC   kidney transplantation or heart damage after cardiac ischemia), treating
CC   epilepsy, specifically generalized epileptic attacks (e.g. petit mal and
CC   tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lope
CC   and complex partial attacks), treating microinfarction (e.g. during and

THiS PAGE BLANK (USPTO)

Mon Sep 9 08:38:36 2002

OM protein - protein search, using sw model

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.
```

Run on: August 29, 2002, 07:58:21 ; Search time 46.79 Seconds
                                            (without alignments)
                                            275.630 Million cell updates/sec

Title: US-09-701-586B-10
Perfect score: 2789
Sequence: 1 MAPKRKASVQTEGSKKQRQG.........EYLIYKESQCRLRYLLEIHL 528

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
                1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Match Length	DB	ID	Description
1	707.5	25.4	1013	4	US-08-860-886-2	Sequence 2, Appli
2	702.5	25.2	1014	4	US-09-078-347A-3	Sequence 3, Appli
3	143	5.1	1327	4	US-09-196-387-2	Sequence 2, Appli
4	106.5	3.8	635	4	US-09-046-992-2	Sequence 2, Appli
5	104	3.7	2285	4	US-09-308-375-2	Sequence 1, Appli
6	103.5	3.7	613	1	US-08-405-615-1	Sequence 1, Appli
7	103.5	3.7	613	2	US-08-461-234-1	Sequence 1, Appli
8	103.5	3.7	613	1	US-08-463-480-1	Sequence 1, Appli
9	103.5	3.7	614	2	US-08-225-224-1	Sequence 1, Appli
10	103.5	3.7	614	1	US-08-722-258-1	Sequence 1, Appli
11	103.5	3.7	614	5	PCT-US95-04468-1	Sequence 1, Appli
12	103	3.7	638	3	US-09-047-148-2	Sequence 2, Appli
13	103	3.7	462	3	US-08-484-438-42	Sequence 42, Appl
14	102.5	3.7	3075	2	US-08-460-309-5	Sequence 5, Appli
15	102.5	3.7	3075	2	US-08-125-077-5	Sequence 5, Appli
16	101	3.6	622	1	US-08-356-786-16	Sequence 16, Appl
17	100.5	3.6	420	1	US-08-391-259-7	Sequence 7, Appli
18	100.5	3.6	420	2	US-08-839-425-7	Sequence 7, Appli
19	100.5	3.6	2482	1	US-08-328-254-6	Sequence 6, Appli
20	99	3.5	579	1	US-08-126-564A-31	Sequence 31, Appl
21	99	3.5	579	5	PCT-US94-09143-31	Sequence 31, Appl
22	98	3.5	889	4	US-09-336-447A-15	Sequence 15, Appl
23	97	3.5	662	1	US-08-224-657-88	Sequence 88, Appl
24	97	3.5	662	4	US-09-354-138-88	Sequence 88, Appl
25	96.5	3.5	497	2	US-08-511-485-4	Sequence 4, Appli
26	96.5	3.5	497	2	US-09-212-971-4	Sequence 4, Appli
27	96.5	3.5	497	4	US-08-800-929A-4	Sequence 4, Appli
28	96.5	3.5	497	4	US-09-617-053A-4	Sequence 4, Appli
29	96	3.4	163	1	US-08-044-618-6	Sequence 6, Appli
30	96	3.4	637	2	US-08-235-838-14	Sequence 14, Appl
31	96	3.4	637	2	US-08-465-473B-14	Sequence 14, Appl
32	95.5	3.4	396	4	US-09-046-992-4	Sequence 4, Appli
33	95.5	3.4	420	1	US-08-391-259-2	Sequence 2, Appli
34	95.5	3.4	420	2	US-08-391-259-11	Sequence 11, Appl
35	95.5	3.4	420	2	US-08-839-425-2	Sequence 2, Appli
36	95.5	3.4	420	2	US-08-839-425-11	Sequence 11, Appl
37	95.5	3.4	599	1	US-08-463-163-3	Sequence 3, Appli
38	95.5	3.4	637	1	US-08-235-838-16	Sequence 16, Appl
39	95.5	3.4	637	4	US-08-465-473B-16	Sequence 16, Appl
40	95.5	3.4	1128	4	US-08-923-992A-6	Sequence 6, Appli
41	95.5	3.4	1164	4	US-08-923-992A-2	Sequence 2, Appli
42	95.5	3.4	3248	1	US-08-353-700-1	Sequence 1, Appli
43	95.5	3.4	3248	5	PCT-US95-16216-1	Sequence 1, Appli
44	95	3.4	920	2	US-08-451-715A-2	Sequence 2, Appli
45	94.5	3.4	1098	4	US-08-923-992A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-860-886-2
; Sequence 2, Application US/08860886
; Patent No. 635009
; GENERAL INFORMATION:
; APPLICANT: Burkle, Alexander
; APPLICANT: Zur Hausen, Harald
; APPLICANT: Jan-Heiner, Kupper
; TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE
; TITLE OF INVENTION: IN GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Pennie & Edmonds, LLP
;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 10036-2811
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: Windows
;   SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/860,886
;   FILING DATE: 03-OCT-1997
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Coruzzi, Laura A
;   REGISTRATION NUMBER: 30,742
;   REFERENCE/DOCKET NUMBER: 8484-0028-999
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650-493-4935
;   TELEFAX: 650-493-5556
;   TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO; 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1013 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-860-886-2

    Query Match   25.4%; Score 707.5; DB 4; Length 1013;

; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1014 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-078-347A-3

```
Query Match     25.2%;  Score 702.5;  DB 4;  Length 1014;
Best Local Similarity  33.8%;  Pred. No. 1e-59;
Matches  182;  Conservative 102;  Mismatches 217;  Indels 217;  Gaps 17;

Qy    1 MAPKRRASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          ::|: ||: :|  :| :   |    ||  ::| |: ::|  | !    :|  :  |
Db  493 VAPRGKSGA---ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 549

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          ::|: :| :: ||   : : :|||||:|   ::| | || ||| :  |:|: ::| ::
Db  550 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKED 609

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVK 175
          | :|: |: ||: :|  | |   | :  |  |: :|  :| :::   |: |  | || |
Db  610 AIEQFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKLTVNPGTKSKLPK 667

Qy  176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 235
          |           :|:|| : :|::||| ||::|||||||||:|:|:  ||:   :  ||
Db  668 P-----VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQAVS 721

Qy  236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
          :   : :|: |::  ||::||| ||  :||:::|  |||| :::| |: |::::| : |
Db  722 QGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLRG 780

Qy  296 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--H 352
          :  :||  :   |: |  |  : |  :   |::|  ||: : |||   :|  |   :
Db  781 GSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATTHSAYDLEVID 835

Qy  353 VWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGI 408
          : :: |||| |  | | || :|||||| : : : |:|:|||||   |     || |||
Db  836 IFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGI 895

Qy  409 YFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 468
          |||  :||||:: |   |  :  ::| | | ||| :::||   ::|  |  | |  || |
Db  896 YFADMVSKSANYYHTSQ--GDPIGLILLGEVALGNMYELK-HASHISRLPKGKHSVKGLG 952

Qy  469 QTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI 526
          :|||||: :|:|:||| ||| |  | ||: || :   ::|: :::  : :|:|||:|
Db  953 KTTPDPSANI--SLDG--VDVPLGTGIS-SGVIDTSLLYNEYIVYDIAQVNLKYLLKL 1005
```

RESULT 3
US-09-196-387-2
; Sequence 2, Application US/0196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
Best Local Similarity  34.0%;  Pred. No. 3.3e-60;
Matches  183;  Conservative 101;  Mismatches 216;  Indels 39;  Gaps 17;

Qy    1 MAPKRRASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
          ::|: ||: :|  :| :   |    ||  ::| |: ::|  |  !   :|  :  |
Db  492 VAPRGKSGA--ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 548

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
          ::|: :| :: ||   : : :|||||:|   ::| | || ||| :  |:|: ::| ::
Db  549 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKED 608

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK--VDSGPVRTVVK 175
          | :|: |: ||: :|  | |   | :  |  |: :|  :| :::   |: |  | || |
Db  609 AIEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKLTVNPGTKSKLPK 666

Qy  176 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 235
          |           :|:|| : :|::||| ||::|||||||||:|:|:  ||:   :  ||
Db  667 P------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQAVS 720

Qy  236 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 295
          :   : :|: |::  ||::||| ||  :||:::|  |||| :::| |: |::::| : |
Db  721 QGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLRG 779

Qy  296 PGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLR 351
          :  :||  :   |: |  |  : |  :   |::|  ||: : |||    :|  |
Db  780 GSDDSSK-----DPIDVNYEKLKTDIKVDRDSEEAEIIRKYVKNTHATTHNAYDL-EVI 833

Qy  352 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKG 407
          : :: |||| |  | | || :|||||| : : : |:|:|||||   |     || |||
Db  834 DIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKG 893

Qy  408 IYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARG 467
          |||  :||||:: |   |  :  ::| | | ||| :::||   ::|  |  | |  || |
Db  894 IYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISRLPKGKHSVKGL 950

Qy  468 GQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI 526
          :|||||: :|:|:||| ||| |  | ||: || :   ::|: :::  : :|:|||:|
Db  951 GKTTPDPSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEYIVYDIAQVNLKYLLKL 1004
```

RESULT 2
US-09-078-347A-3
; Sequence 3, Application US/09078347A
; Patent No. 6132968
; GENERAL INFORMATION:
; APPLICANT: Le, Xiao-Chun
; APPLICANT: Weinfeld, Michael
; APPLICANT: Xing, James Z.
; TITLE OF INVENTION: Methods for Quantitating Low Level
; TITLE OF INVENTION: Modifications of Nucleotide Sequences
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,347A
; FILING DATE: 13-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UALB-03283
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/196,387
;   FILING DATE:
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/095,225
;   FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Jackson Esq., David A.
;   REGISTRATION NUMBER: 26,742
;   REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMUNICATION INFORMATION:
;   TELEPHONE: 201-487-5800
;   TELEFAX: 201-343-1684
;   TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1327 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
US-09-196-387-2

Query Match   5.1%; Score 143; DB 4; Length 1327;
Best Local Similarity 23.9%; Pred. No. 9.5e-05;
  Matches 71; Conservative 41; Mismatches 107; Indels 78; Gaps 13;

Qy   271 DVLQAKKDMLLVLADI---ELAQTLQAAPGEEEKVEEVPHPL------DRDYQLLRCQLQ 322
              |:| | |||:| || : | | | | : ||| ||:||||:||:
Db   1045 DIFETEQITLDVLADMGHEELKEIGINAYGHRHKLLIKGVERLLGGQGTNPYLTFHCVNQ 1104

Qy   323 ---LLDSG--ESEYKAIQTYLKQT-----------GNSYRCPNLRHVWKVNREGEGDRF-- 365
             :||| |: |: ||||:|:| :|: | |||: ||:||||||
Db   1105 GTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCH 1164

Qy   366 ------QAHSKLGNRRLLWHGTNVAVVAAILTSGL-RIMPHSGGRVGKGIYFASENSKSA 418
              | |||:|:: |:|| :| || : |||||| ::||| |||||||||||| :|||
Db   1165 RQKEVSEENHHNHNERMLFHGS--PFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSN 1222

Qy   419 GYVTTM----HCGGHQ------VGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIAR 467
          ||:| | |||| | : |: || | : |||:||| ||:|||||||||||
Db   1223 QYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKS-FLQFSTMKMAHAPPGHHSVI-- 1279

Qy   468 GQTEPDPAQDIELELDGQPVVVPQGGPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLL 524
              :|:|: | | | :: : | :| :||:| ||: ||:::||:
Db   1280 --------------------GRPSV-------------------NGLAYAEYVIYRGEQAYPEYLI 1306

RESULT 4
US-09-046-992-2
; Sequence 2, Application US/09046992
; Patent No. 6140066
; GENERAL INFORMATION:
;   APPLICANT: Lorberboum-Galski, Haya
;   APPLICANT: Yarkoni, Shai
;   APPLICANT: Ben-Yehudah, Ahmi
;   TITLE OF INVENTION: METHODS OF CANCER DIAGNOSIS
;   TITLE OF INVENTION: USING A CHIMERIC TOXIN
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Pennie & Edmonds, LLP
;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 10036-2811
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/046,992
;   FILING DATE: 24-MAR-1998
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Poissant, Brian M
;   REGISTRATION NUMBER: 28,462
;   REFERENCE/DOCKET NUMBER: 9457-0013-999
; TELECOMUNICATION INFORMATION:
;   TELEPHONE: 650-493-4935
;   TELEFAX: 650-493-5556
;   TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 635 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   FRAGMENT TYPE: internal
US-09-046-992-2

Query Match   3.8%; Score 106.5; DB 4; Length 635;
Best Local Similarity 19.0%; Pred. No. 0.11;
  Matches 79; Conservative 55; Mismatches 154; Indels 127; Gaps 17;

Qy   148 TLIEVQGEAESQEAVVKVDSGPVRTVVKP------------CSLDPATQNLITNIFSKEM 195
              ::: :| | : | | || | | | : || :||| |: ||
Db   195 SVVMAQNQPRREKRWSEWASGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNP 254

Qy   196 FKNAMTLMNLDVK------------------------KMPLGKLTKQQIARGFEALE-- 228
          |: :   ::| :| ||| | : :: ||
Db   255 AKH------DLDIKPTVISEELEFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQC 309

Qy   229 ------------------------ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPP 265
                                     ||||||||| || : :: |: : |:|||| | ::
Db   310 GYPVQRLVLALYLAARLSWNQVDQVIRNALASP-GSGGDIGEAI--------------REQPE 356

Qy   266 PINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL-- 323
              | | :| | | |: | : | | | | : ||| |||:||:||
Db   357 ------------QAR-LALTLAAAESRFVRQGTGNDEAGA-----ANADVVSLTCPVAAGE 400

Qy   324 ---LDSG----ESEYKAIQTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRR 375
             :||| | |: ||||:| |: |||:|||:||: ||:|| |:| :| :
Db   401 CAGPADSGDALLEANYPTGAEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERG 455

Qy   376 LL--WHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVG 432
           | | | | : |:| |: || |||||||||||||||| :|||||| |||| ||
Db   456 YVFVGYHGTFLEAAQSIVFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR--- 512

Qy   433 YMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 487
          : | : ||| | : : ||::|||||||:| | || | ||:|| |:
Db   513 ----GRIRNGALLRVYVPRSSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 557

RESULT 5
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
;   APPLICANT: Genencor International, Inc.
;   TITLE OF INVENTION: Proteases From Gram-Positive Organisms
;   FILE REFERENCE: GC394-PCT
;   CURRENT APPLICATION NUMBER: US/09/308,375
;   CURRENT FILING DATE: 1999-05-14
;   EARLIER APPLICATION NUMBER: EP9719636.4
;   EARLIER FILING DATE: 1997-09-15

us-09-701-586b-10.rai

```
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 2285
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-308-375-2

Query Match        3.7%;      Score 104;  DB 4;  Length 2285;
Best Local Similarity  21.2%;  Pred. No. 1.5;
Matches  46;  Conservative  47;  Mismatches  94;  Indels  30;  Gaps  7;

Qy  116 EDAKKDFKK----------FWEKTKNKW-----EERDRFVAQPNKYTLIEVQGEAES 158
          ||||||                ::     ||   :: |  | : |
Db  881 EDASKEIKKSKDELKQYKQIADYNDKGRPKWDLIADDDDYKVAADKAKQSMLKAQSDIES 940
Qy  159 QEAVVKVDSGPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQ 218
          | ||          ::  |||| |:|| | :|: || |  :|||| || |: |||
Db  941 GNAKVKDSVLSIANAYSSIDISNTLKTSISDVVNKLNLKDDLDPEELEKFSSSLGKL-QE 999
Qy  219 QIARGFEALEALEAMKNPTGDGQSLEEL--SSCFYTVIPHNFGRSRPPPINSPDVLQA 275
          :: :: ::  | | ||| |||   :   :   || || || :
Db  1000 KMQKALDSGD--EKAFDNAKKDLQSLLEYSKSDSSIDVFKMSFDKAQ-----KNIKDG 1051
Qy  276 KKDMLLVLADI-ELAQTLQAAPGEEEKVEEVPHPLD 311
          ||     | | || :|  ||  ||
Db  1052 DKSLSSVKSEVGDLGETLAEAGNEAEDFGKKLKEALD 1088

RESULT 6
US-08-405-615-1
; Sequence 1, Application US/08405615
; Patent No. 5602095
; GENERAL INFORMATION:
;   APPLICANT: Pastan, Ira
;   APPLICANT: FitzGerald, David J.
;   TITLE OF INVENTION: Recombinant Pseudomonas Exotoxin with
;   TITLE OF INVENTION: Increased Activity
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Ellen L. Weber
;     STREET: One Market Plaza, Steuart Tower, Suite 2000
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94105
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/405,615
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/07/901,709
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Weber, Ellen L.
;     REGISTRATION NUMBER: 32,762
;     REFERENCE/DOCKET NUMBER: 15280-36
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-543-9600
;     TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 613 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
```

```
US-08-405-615-1

Query Match        3.7%;      Score 103.5;  DB 1;  Length 613;
Best Local Similarity  19.7%;  Pred. No. 0.21;
Matches  78;  Conservative  47;  Mismatches  144;  Indels  127;  Gaps  17;

Qy  167 SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 208
          :||                   ||  |:  :| | :::   :||::
Db  192 SGKVLCLLDPLDGVNYLAQRCNLDDTWEGKIYRVLAGNPAKH----DLDIKPTVISH 246
Qy  209 ---------KMPLGKLTKQQIARGFEALE 228
                  |:||| ||| |  |:| | ||
Db  247 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 306
Qy  229 ---ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 284
          :: |: |  | :|: |       :: : | ||
Db  307 QVDQVIRNALASP-GSGGDLGEAI-----REQPE--------QAR--LALTLA 343
Qy  285 DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL----LDSG-----ESEYKAI 334
          | |:    :: | |  ||||:| | :  |:     : |  ||:     :: |
Db  344 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 397
Qy  335 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL--WHGTNVAVVAAILT 391
          :|: | |       :|| |: | :::  | |        :|||  ::::: :
Db  398 AEFLGDGGDVSFSTRGTQNWTVER------LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 452
Qy  392 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTMHCCGHQVGYMFLGEVALGKEHHITIDD 451
          :|| | :  ||     :: |:: ||       ::  ||
Db  453 GGVRARSQDLDAIWRGFYIAGDPALAYGVYAQDQEPDAR-------GRIRNGALLRYVYPR 505
Qy  452 PSLKSPPPGEFDSVIARGQTEPDPAQDIELELDGQPV 487
          : :  || | |   |   |: |   |  : |:| |:|
Db  506 SSL---PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 535
```

```
RESULT 7
US-08-461-234-1
; Sequence 1, Application US/08461234
; Patent No. 5821238
; GENERAL INFORMATION:
;   APPLICANT: Pastan, Ira H.
;   APPLICANT: FitzGerald, David J.
;   TITLE OF INVENTION: Recombinant Pseudomonas Exotoxin with
;   TITLE OF INVENTION: Increased Activity
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew
;     STREET: One Market Plaza, Steuart Street Tower
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94105-1492
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/461,234
;     FILING DATE: 05-JUN-1995
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/901,709
;     FILING DATE: 18-JUN-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/405,615
;     FILING DATE: 15-MAR-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Weber, Ellen Lauver
;     REGISTRATION NUMBER: 32,762
;     REFERENCE/DOCKET NUMBER: 15280-36-3
;   TELECOMMUNICATION INFORMATION:
```

```
;     TELEPHONE: (415) 543-9600
;     TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 613 amino acids
;     TYPE: amino acid
;     STRANDEDNESS:
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
US-08-461-234-1

Query Match     3.7%;     Score 103.5;  DB 2;  Length 613;
Best Local Similarity  19.7%;  Pred. No. 0.21;
Matches  78;  Conservative  47;  Mismatches 144;  Indels 127;  Gaps  17;

Qy  167  SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 208
              |||  ::|         |:||  |  ||  ||:|
Db  192  SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 246

Qy  209  -----------KMPLGKLTKQQIARGFEALE----------------- 228
                        ||| |: :  ||:| ||
Db  247  RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 306

Qy  229  ----ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 284
             :|| |::: ||  |
Db  307  QVDQVIRNALASP-GSGGDLGEAI--------REQPE-------QAR--LALTLA 343

Qy  285  DIELAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAI 334
              ||                 |:     ||:  |||          ||       |
Db  344  AAESERFVRQGTGNDEAGA-----ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 397

Qy  335  QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL--WHGTNVAVVAAILT 391
              :|             ||| |      :||| |   ::|:
Db  398  AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 452

Qy  392  SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 451
              :|:           ||| |       ||  :
Db  453  GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 505

Qy  452  PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 487
              |||         ||:: |:|  |
Db  506  SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 535

RESULT  8
US-08-463-480-1
; Sequence 1, Application US/08463480
; Patent No. 5854044
; GENERAL INFORMATION:
;   APPLICANT: Pastan, Ira H.
;   APPLICANT: FitzGerald, David J.
;   TITLE OF INVENTION: Recombinant Pseudomonas Exotoxin with
;                       Increased Activity
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, Eighth Floor
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/463,480
;     FILING DATE: 05-JUN-1995
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
```

```
;   APPLICATION NUMBER: US 07/901,709
;   FILING DATE: 18-JUN-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/405,615
;     FILING DATE: 15-MAR-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Weber, Ellen Lauver
;     REGISTRATION NUMBER: 32,762
;     REFERENCE/DOCKET NUMBER: 15280-36-2
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 576-0200
;     TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 613 amino acids
;     TYPE: amino acid
;     STRANDEDNESS:
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
US-08-463-480-1

Query Match     3.7%;     Score 103.5;  DB 2;  Length 613;
Best Local Similarity  19.7%;  Pred. No. 0.21;
Matches  78;  Conservative  47;  Mismatches 144;  Indels 127;  Gaps  17;

Qy  167  SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 208
              |||  ::|         |:||  |  ||  ||:|
Db  192  SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 246

Qy  209  -----------KMPLGKLTKQQIARGFEALE----------------- 228
                        ||| |: :  ||:| ||
Db  247  RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 306

Qy  229  ----ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 284
             :|| |::: ||  |
Db  307  QVDQVIRNALASP-GSGGDLGEAI--------REQPE-------QAR--LALTLA 343

Qy  285  DIELAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAI 334
              ||                 |:     ||:  |||          ||       |
Db  344  AAESERFVRQGTGNDEAGA-----ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 397

Qy  335  QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL--WHGTNVAVVAAILT 391
              :|             ||| |      :||| |   ::|:
Db  398  AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 452

Qy  392  SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 451
              :|:           ||| |       ||  :
Db  453  GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 505

Qy  452  PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 487
              |||         ||:: |:|  |
Db  506  SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 535

RESULT  9
US-08-225-224-1
; Sequence 1, Application US/08225224
; Patent No. 5635599
; GENERAL INFORMATION:
;   APPLICANT: KREITMAN, Robert J.
;   APPLICANT: PASTAN, Ira
;   TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
;   TITLE OF INVENTION: CIRCULARLY PERMUTED FUSION PROTEINS
;   NUMBER OF SEQUENCES: 57
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend Khourie and Crew
;     STREET: Steuart Street Tower, One Market Plaza
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: US
;     ZIP: 94105-1493
```

```
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/225,224
    FILING DATE: 8-APR-1994
    CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
    NAME: Weber, Ellen L.
    REGISTRATION NUMBER: 32,762
    REFERENCE/DOCKET NUMBER: 15280-193
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (415) 543-9600
    TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
    LENGTH: 614 amino acids
    TYPE: amino acid
    STRANDEDNESS: unknown
    TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
    NAME/KEY: Protein
    LOCATION: 1..614
    OTHER INFORMATION: /label= native-PE
US-08-225-224-1

Query Match             3.7%;   Score 103.5; DB 1; Length 614;
Best Local Similarity  19.7%;   Pred. No. 0.21;
Matches    78;  Conservative    47;  Mismatches  144;  Indels 127;  Gaps  17;

Qy  167 SGPVRTVVKP-----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 208
            |||  ::|     |:||   |  ::   |:||
Db  193 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 247

Qy  209 ------------------------KMPLGKLTKQQIARGFEALE--------------- 228
                                ||:|: ||:||||
Db  248 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 307

Qy  229 ---ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 284
          :|:|  ||||                                 ||   ||||
Db  308 QVDQVIRNALASP-GSGGDLGEAI------------------REQPE--------QAR--LALTLA 344

Qy  285 DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAI 334
          ||:   |:| ||                               |||     |||:|
Db  345 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 398

Qy  335 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL--WHGTNVAVVAAILT 391
          :| :   || |   | ||||           ||:|  |:
Db  399 AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 453

Qy  392 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 451
          |:|            :|:|:|: :  ||
Db  454 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVVYVPR 506

Qy  452 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 487
          |:|       |||      |
Db  507 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 536


RESULT 10
US-08-722-258-1
; Sequence 1, Application US/08722258
; Patent No. 6011002
; GENERAL INFORMATION:
APPLICANT: Pastan, Ira
APPLICANT: Kreitman, Robert J.
APPLICANT: Puri, Raj K.
TITLE OF INVENTION: Circularly Permuted Ligands and
TITLE OF INVENTION: Circularly Permuted Chimeric Molecules
```

```
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Townsend and Townsend and Crew LLP
    STREET: Two Embarcadero Center, Eighth Floor
    CITY: San Francisco
    STATE: California
    COUNTRY: USA
    ZIP: 94111-3834
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/722,258
    FILING DATE: 08-JAN-1997
    CLASSIFICATION: 530
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: WO PCT/US95/04468
    FILING DATE: 06-APR-1995
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/225,224
    FILING DATE: 08-APR-1994
ATTORNEY/AGENT INFORMATION:
    NAME: Weber, Ellen Lauver
    REGISTRATION NUMBER: 32,762
    REFERENCE/DOCKET NUMBER: 015280-19310OUS
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (415) 576-0200
    TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
    LENGTH: 614 amino acids
    TYPE: amino acid
    STRANDEDNESS:
    TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
    NAME/KEY: Protein
    LOCATION: 1..614
    OTHER INFORMATION: /note= "native Pseudomonas exotoxin
    OTHER INFORMATION:
US-08-722-258-1

Query Match             3.7%;   Score 103.5; DB 3; Length 614;
Best Local Similarity  19.7%;   Pred. No. 0.21;
Matches    78;  Conservative    47;  Mismatches  144;  Indels 127;  Gaps  17;

Qy  167 SGPVRTVVKP-----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 208
            |||  ::|     |:||   |  ::   |:||
Db  193 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 247

Qy  209 ------------------------KMPLGKLTKQQIARGFEALE--------------- 228
                                ||:|: ||:||||
Db  248 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 307

Qy  229 ---ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 284
          :|:|  ||||                                 ||   ||||
Db  308 QVDQVIRNALASP-GSGGDLGEAI------------------REQPE--------QAR--LALTLA 344

Qy  285 DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAI 334
          ||:   |:| ||                               |||     |||:|
Db  345 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 398

Qy  335 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL--WHGTNVAVVAAILT 391
          :| :   || |   | ||||           ||:|  |:
Db  399 AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 453

Qy  392 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 451
          |:|            :|:|:|: :  ||
Db  454 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVVYVPR 506
```

```
Qy    452  PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV  487
           ||    |||   |:|:|  |  ||||
Db    507  SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL  536
```

RESULT 11
PCT-US95-04468-1
; Sequence 1, Application PC/TUS9504468
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
; TITLE OF INVENTION: CIRCULARLY PERMUTED FUSION PROTEINS
; NUMBER OF SEQUENCES: 59
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04468
; FILING DATE: 07-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,224
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-193-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 614 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1.614
; OTHER INFORMATION: /label= native-PE
PCT-US95-04468-1

```
Query Match       3.7%;  Score 103.5;  DB 5;  Length 614;
Best Local Similarity  19.7%;  Pred. No. 0.21;
Matches   78;  Conservative   47;  Mismatches 144;  Indels 127;  Gaps   17;

Qy  167  SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------  208
         || ||: :||              :  ||  :    | :    :||:|
Db  193  SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH  247
Qy  209  -------------------KMPLGKLTKQQIARGFEALE  228
                            ||  |:|| ||:|  ||
Db  248  RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN  307
Qy  229  ---ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA  284
           :   |:   ||
Db  308  QVDQVIRNALASP-GSGGDLGEAI---------REQPE-------QAR--LALTLA  344
Qy  285  DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL------LDSG-----ESEYKAI  334
         |  | |:|    ||       |    ||  |:|       ||         |
Db  345  AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG  398
Qy  335  QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL--WHGTNVAVVAAILT  391
         ::  |     |           |    | :|                |::
Db  399  AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYFVFVGYHGTFLEAAQSIVF  453
Qy  392  SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDD  451
         :: | |   || :  ||  |     :  |      :    |  :  :
Db  454  GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR  506
```

```
Qy    452  PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV  487
           ||    |||   |:|:|  |  ||||
Db    507  SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL  536
```

RESULT 12
US-09-047-148-2
; Sequence 2, Application US/09047148
; Patent No. 6086900
; GENERAL INFORMATION:
; APPLICANT: Draper, Rockford
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USING
; TITLE OF INVENTION: MEMBRANE-PENETRATING PROTEINS TO CARRY MATERIALS ACROSS
; TITLE OF INVENTION: CELL MEMBRANES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/047,148
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/042,056
; FILING DATE: 26-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37, 642
; REFERENCE/DOCKET NUMBER: UTSF:072
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
US-09-047-148-2

```
Query Match       3.7%;  Score 103.5;  DB 3;  Length 638;
Best Local Similarity  19.7%;  Pred. No. 0.22;
Matches   78;  Conservative   47;  Mismatches 144;  Indels 127;  Gaps   17;

Qy  167  SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------  208
         || ||: :||              :  ||  :    | :    :||:|
Db  217  SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH  271
Qy  209  -------------------KMPLGKLITKQQIARGFEALE  228
                            ||  |:|| ||:|  ||
Db  272  RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN  331
Qy  229  ----ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA  284
           :   |:   ||
Db  332  QVDQVIRNALASP-GSGGDLGEAI---------REQPE-------QAR--LALTLA  368
Qy  285  DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL------LDSG-----ESEYKAI  334
         |  | |:|    ||       |    ||  |:|       ||         |
Db  369  AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG  422
Qy  335  QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL---WHGTNVAVVAAILT  391
         ::  |     |           |    | :|                |::
```

```
Db    22 KKKGGKNGKNR-RNRSHLIKCAEKEKTFCVNG---GECFTVKDLSNPSRYLCK-CPNEFT 76
                  |||      |:|       ||            ||||   |||  |:|   |:|| ||
Qy   181 -PATQNLITNIFSK--EMFK------NAMTLMNLDVKKMPLGKLTKQQTARGFEALE-- 228
                  :||   | |:|             ||:|     | |||:|| :||   |:||
Db    77 GDRCRQNYVMASFYKAEELYKLMAEEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQC 136
Qy   229 ------------ALEEAMKNPTGDQSLEELSSCFYTVIPHNFGRSRPP 265
                          :||:      ||| || :|  |: :| |  |||:|
Db   137 GYPVQRLVALYLAARLSWNQVDQVIRNALASP-GSGGDLGEAI---------REQPE 183
Qy   266 PINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQL-- 323
                  |:  ||      |:|    ::|
Db   184 ------QAR--LALTLAAAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGE 227
Qy   324 ----LDSG---ESEYKAIQTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRR 375
                  |||    |  | :|  :|  ||      |||  ||
Db   228 CAGPADSGDALLERNYPTGAEFLGDGDVSFSRGTQNWTVER----LLQAHRQLEERG 282
Qy   376 LL--WHGTNVAVVAILTSGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVG 432
                  |       :||| :|     | |||||:| ||||||||||||||||||||||||||
Db   283 YVFVGYHGTFLEAAQSIVFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR--- 339
Qy   433 YMFLGEVALGKEHHITIDDPSLKSPPPGFGFDSVI------ARGQTEPDPAQDIELELD-- 483
                  | ||: |||:||||||| |||||||||| |||       ||||||||||||:||| |
Db   340 ---GRIRNGALLRVYVPRSSL----PGFYRTSLTLAGGEAAGEVERLIGHPLPLRLDAI 391
Qy   484 ------------GQP----VVVPQGPP 494
                              ||:     ||::|
Db   392 TGPEEEGGRLETILGWPLAERTVVIPSAIP 421
```

RESULT 14
US-08-460-309-5
; Sequence 5, Application US/08460309
; Patent No. 5837496
; GENERAL INFORMATION:
;   APPLICANT: Engvall, Eva
;   APPLICANT: Leivo, Ilmo
;   TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
;   TITLE OF INVENTION: Fragments and Uses Thereof
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Campbell and Flores
;   STREET: 4370 La Jolla Village Drive, Suite 700
;   CITY: San Diego
;   STATE: California
;   COUNTRY: USA
;   ZIP: 92122
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/460,309
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/125,077
;   FILING DATE: 22-SEP-1993
;   APPLICATION NUMBER: US PCT/US 94/10730
;   FILING DATE: 21-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/472,319
;   FILING DATE: 30-JAN-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/919,951
;   FILING DATE: 27-JUL-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Campbell, Cathryn A.
;   REGISTRATION NUMBER: 31,815
;   REFERENCE/DOCKET NUMBER: P-LA 9721

```
Db   423 AEFLGDGDGVSFSFSRGTQNWTVER-----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 477
                  |:|                      :||   | |:|
Qy   392 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 451
                  ||                       :|  ||             :|           :
Db   478 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 530
                  ||   ||                     |:|      |  |:|
Qy   452 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 487
                  ||       |                 | :|
Db   531 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 560
```

RESULT 13
US-08-484-438-42
; Sequence 42, Application US/08484438
; Patent No. 5811098
; Patent No. 5811098 5780031
; GENERAL INFORMATION:
;   APPLICANT: Plowman, Gregory D.
;   APPLICANT: Culouscou, Jean-Michel
;   APPLICANT: Shoyab, Mohammed
;   APPLICANT: Siegall, Clay B.
;   APPLICANT: Hellstr m, Ingegerd
;   APPLICANT: Hellstr m, Karl E.
;   TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
;   NUMBER OF SEQUENCES: 42
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Pennie & Edmonds
;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 10036-2711
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/484,438
;   FILING DATE: 07-JUN-1995
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/323,442
;   FILING DATE: 14-OCT-1994
;   APPLICATION NUMBER: US 08/150,704
;   FILING DATE: 10-NOV-1993
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/981,165
;   FILING DATE: 24-NOV-1992
;   CLASSIFICATION: 530
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Misrock, S. Leslie
;   REGISTRATION NUMBER: 18,872
;   REFERENCE/DOCKET NUMBER: 5624-230
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 790-9090
;   TELEFAX: (212) 869-8864/9741
;   TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 42:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 462 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-484-438-42

  Query Match   3.7%; Score 103; DB 2; Length 462;
  Best Local Similarity 21.1%; Pred. No. 0.15;
  Matches   95; Conservative 46; Mismatches 181; Indels 128; Gaps 21;

Qy   123 KKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCSLD-- 180
```

```
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/125,077
;   FILING DATE:  22-SEP-1993
;   CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US PCT/US 94/10730
;   FILING DATE:  21-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 07/472,319
;   FILING DATE:  30-JAN-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 07/919,951
;   FILING DATE:  27-JUL-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Campbell, Cathryn A.
;   REGISTRATION NUMBER:  31,815
;   REFERENCE/DOCKET NUMBER:  P-LA 9721
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (619) 535-9001
;   TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  3075 amino acids
;   TYPE:  amino acid
;   TOPOLOGY:  linear
US-08-125-077-5


Query Match          3.7%;   Score 102.5;  DB 2;   Length 3075;
Best Local Similarity  19.7%;   Pred. No. 3.3;
Matches 104;  Conservative  79;  Mismatches 191;  Indels 155;  Gaps  23;

Qy     10 QTEGSKKQR-----QGTEEDSFRST-------AEALRAAPADNRVIRVDPSCPFSRNP  56
             ||   ||   |:   |:      |     :  |: |  ::  |  ||   ::
Db   1500 QTPGGSCQKCDCNPHGSVHGDCDRTSGQCVCRLGASGLRCDECEPRHILMETDCVSCDDE 1559

Qy     57 --GIQVHE-----DYDCTLNQTNI------------GNNNNKFYIIQLLEEGSRFFCWNRWG  99
            |: |:|       | :||||  :           :||
Db   1560 CVGVLLNDLDEIGDAVLSLNLTGIIPVPYGILSNLENTTKYLQESLLKENMQ---------- 1611

Qy    100 RVGEVGQSKMNHFTCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQ 159
             ::|:|:       |:|     |: |:  :| | |     ||   |  ||
Db   1612 --KDLGKIKLEGVA----------EETDNLQKKLTRMLASTQKVNRATERIFKESQ 1655

Qy    160 EAVVKVDSGPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKK---MPLGKLT 216
            : :::        | :   | |     | |    |:     ::| :   |
Db   1656 DLAVAIE----------RLQMSITEIMEK------TTLNQTLDEDFLLPNSTLQ 1693

Qy    217 KQQIARGFEALEALE-----EAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPD 271
            |::   |  ||  :|     |:   |  ::| || : ::| || |   | : |    |
Db   1694 NMQ-QNGTSLLEIMQIRDFTQLHQNATLELKAAEDLLS---QIQENYQK----PLEELE 1744

Qy    272 VLQAKKDMLLVLADIELAQTLQAAPGEEEEKVEEVPHPL----------DRDYQLLRCQLQ-- 322
            ||:   |: |   ||   |  :     | :|: | : :|          :|  |  :|
Db   1745 VLKEAASHVLSKHNNEL-KAEEALVREAEAKMQESNHLLLMVNANLREFSDKKLHVQEEQ 1803

Qy    323 ------------LLDSGESEYKAIQTYLKQTGNSYRCPNLRH-----VWKVNREGEGDRF 365
                        | :|    :  ::  | |  | | |:|       |:
Db   1804 NLTSELIVQGRGLIDAAAAQTDAVQDALEHLED---------HQDKLLLWSAKIRHHIDDL 1855

Qy    366 QAHSKLGNR-----RLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGY 420
            :|: :|:|:     |  |  : | |  :| ||: | |: |  | |: |::| | | | :
Db   1856 VMHMSQRNAVDLVYRAEDHATEFQRLADVLYSGL---------ENIRNVSLNATSAAY 1904

Qy    421 VTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQ 469
            |                 |  | |: |:|  |: : :    |:| |:
Db   1905 V-----HYNIQSLIEESEELARDAHRTVTETSLLS----ESLVSNGK 1942
```

Search completed: August 29, 2002, 07:58:25

us-09-701-586b-10.rai

Mon Sep 9 08:38:37 2002

Job time: 302 sec

```
                GenCore version 4.5
        Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 29, 2002, 08:02:19 ; Search time 36.19 Seconds
                                        (without alignments)
                                        570.255 Million cell updates/sec

Title:             US-09-701-586B-8
Perfect score:     2813
Sequence:          1 MAPKRRASVQTEGSKKQRQG.........EYLIYKESQCRLRYLLEIHL 533

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters:    105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 45 summaries

Database :         SwissProt_40:*

  Pred.  No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2229	79.2	533	1	PPO3_HUMAN	Q9y6f1 homo sapien
2	713.5	25.4	1012	1	PPO3_CRIGR	Q9r152 cricetulus
3	708	25.2	559	1	PPO2_MOUSE	O88554 homo musculu
4	705	25.1	1013	1	PPOL_HUMAN	P09874 homo sapien
5	702	25.0	1015	1	PPOL_BOVIN	P18493 bos taurus
6	696.5	24.8	583	1	PPO2_HUMAN	Q9un5 homo sapien
7	694.5	24.7	1013	1	PPOL_RAT	P27008 rattus norv
8	691.5	24.6	1011	1	PPOL_CHICK	P26446 gallus gall
9	690	24.5	637	1	PPOL_ARATH	Q11207 arabidopsis
10	686.5	24.4	998	1	PPOL_XENLA	P31669 xenopus lae
11	683.5	24.3	1012	1	PPOL_MOUSE	P11103 mus musculu
12	645.5	22.9	996	1	PPOL_SARPE	Q11208 sarcophaga
13	619	22.0	994	1	PPOL_DROME	P35875 drosophila
14	396.5	14.1	538	1	YON4_CAEEL	Q09525 caenorhabdi
15	275	9.8	1724	1	PPOV_HUMAN	Q9ukk3 homo sapien
16	193	6.9	135	1	PPOL_ONCMA	Q08824 oncorhynchu
17	117	4.2	2035	1	EVPL_MOUSE	Q9d952 mus musculu
18	106.5	3.8	638	1	TOXA_PSEAE	P11439 pseudomonas
19	106.5	3.8	1031	1	KINH_STRPU	P35978 strongyloce
20	104.5	3.7	963	1	KINH_HUMAN	P33176 homo sapien
21	104	3.7	1395	1	SP41_YEAST	P38904 saccharomyc
22	104	3.7	3210	1	CENF_HUMAN	P49454 homo sapien
23	102	3.6	752	1	CAT1_MYCFO	O08404 mycobacteri
24	101.5	3.6	1901	1	YCF1_TOBAC	P12222 nicotiana t
25	101	3.6	1085	1	RBP2_BOVIN	P48820 bos taurus
26	101	3.6	3075	1	LMA1_HUMAN	P25391 homo sapien
27	100.5	3.6	1574	1	RPOC_AQUAE	O67763 aquifex aeo
28	100	3.6	583	1	LAM1_XENLA	P09010 xenopus lae
29	99	3.5	495	1	NUSA_ECOLI	P03003 escherichia
30	99	3.5	497	1	BIR4_HUMAN	P98170 homo sapien
31	99	3.5	638	1	RPA1_EUPOC	P28363 euplotes oc
32	99	3.5	646	1	YE14_SCHPO	O13869 schizosacch
33	99	3.5	890	1	RB6K_HUMAN	O95235 homo sapien
34	98.5	3.5	583	1	T2F1_FLAOK	P14870 flavobacter
35	98.5	3.5	687	1	NNP1_DROME	Q9vjz7 drosophila
36	98	3.5	612	1	EXO2_BPT5	P11109 bacteriopha
37	98	3.5	830	1	YBMA_SCHPO	Q10332 schizosacch
38	98	3.5	963	1	KINH_MOUSE	Q61768 mus musculu
39	97.5	3.5	379	1	YO11_MOUSE	P11260 mus musculu
40	97.5	3.5	603	1	US26_HCMVA	P09699 human cytom
41	97	3.4	725	1	ADDB_MOUSE	Q9qyb8 mus musculu
42	96.5	3.4	1005	1	E4L2_HUMAN	O43491 homo sapien
43	96.5	3.4	2453	1	NCR1_MOUSE	Q60974 mus musculu
44	96.5	3.4	5327	1	ACF7_MOUSE	Q9qxz0 mus musculu
45	96	3.4	629	1	PAB2_ARATH	P42731 arabidopsis

ALIGNMENTS

```
RESULT 1
PPO3_HUMAN
ID   PPO3_HUMAN    STANDARD;    PRT;    533 AA.
AC   Q9Y6F1; Q9UG81;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+) ADP-
DE   ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (pADPRT-3)
DE   (hPARP-3).
GN   ADPRTL3 OR PARP3 OR ADPRT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=9263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [2]
RP   SEQUENCE OF 75-533 FROM N.A.
RC   TISSUE=kidney;
RA   Ansorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels are in
CC       the kidney, skeletal muscle, liver, heart and spleen; also
CC       detected in pancreas, lung, placenta, brain, leukocytes, colon,
CC       small intestine, ovary, testis, prostate and thymus.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics  and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF083068; AAD29855.1; -.
DR   EMBL; AL050034; CAB43246.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation.
FT   DOMAIN       14     20       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   CONFLICT     80     80       K -> N (IN REF. 2).
```

```
FT   CONFLICT        171    171        A -> G (IN REF. 2).
FT   CONFLICT        411    411        K -> E (IN REF. 2).
SQ   SEQUENCE        533 AA;  60117 MW;  7COAB89E64D1B9FD CRC64;

Query Match                    79.2%;  Score 2229;  DB 1;  Length 533;
Best Local Similarity          79.8%;  Pred. No. 1.2e-154;
Matches   430;  Conservative   37;  Mismatches   60;  Indels   12;  Gaps    5;

Qy    1  MAPKRKASVQTEG--SKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI   58
         ||||    ||||     |  |  ||| |||:||::||  |:||||:||  | |||
Db    1  MAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT   60

Qy   59  QVHEDYDCTLNQTIGNNNNKFYIIQLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED  117
         ||:|| |   |||| |||||||||||| | ||| || |||| ||||:||||||||||:|
Db   61  QVYEDYNCTLNQTIENNNKKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED  120

Qy  118  AKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT  177
         |||:|||||||:| || ||||: |  |||||||| |:|:|||  |:|     || ||||
Db  121  AKDFEKKFPREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVK-----VDRAPVRT  175

Qy  178  V---VKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA  234
           |:|||||||||||  || |||||| :|  |:|:|||||||||||||||||||||||||
Db  176  VTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA  235

Qy  235  LEFAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELA  294
         |||  |||||||||||| || ||||||||||  |||||||||||:|||||||||||||||
Db  236  LEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELIQAKKDMLLVLADIELA  295

Qy  295  QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPN  354
         |||  |    ||::|||||||||||||||||||||||| ||||:| |::|:||||||| |
Db  296  QALQAV-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT  354

Qy  355  LRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIY  414
         |||:|||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  355  LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIY  414

Qy  415  FASENSKSAGYVTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQ  474
         |||||||||||  | | |||||||||||| || ||| |:||| |||||||||||| |
Db  415  FASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH  474

Qy  475  TEPPPAQDIELELDGQPVVVPQGPVVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL  533
         ||||  |:|||||||||||||||||||||:|||||||||||||||||||||||||| |
Db  475  TEPDPTQDTELELDGQOVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL  533

RESULT  2
PPOL_CRIGR
ID   PPOL_CRIGR      STANDARD;      PRT;   1012 AA.
AC   Q9R152;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21276334; PubMed=11382339;
RA   Ganesh A., Phillips E., Thacker J., Meuth M.;
RT   "Suppression of the radiation-sensitive phenotype of hamster irs1 and
RT   irs2 strains selected for resistance to 3'-aminobenzamide.";
RL   Int. J. Radiat. Biol. 77:609-616(2001).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
```

```
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE (BY
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor =
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute.  There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF168781; AAD45817.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET       0      0                                BY SIMLARITY.
FT   DNA_BIND       1    371                                BY SIMILARITY.
FT   DOMAIN       372    522        AUTOMODIFICATION DOMAIN (BY SIMILARITY).
FT   DOMAIN       386    475        BRCT.
FT   DOMAIN       523   1012        NAD-BINDING (BY SIMILARITY).
FT   ZN_FING       20     55        PARP-TYPE.
FT   ZN_FING      124    161        PARP-TYPE.
FT   DOMAIN       206    208        NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                  (POTENTIAL).
FT   DOMAIN       220    225        NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                  (POTENTIAL).
FT   MOD_RES      406    406        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      412    412        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      434    434        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      436    436        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      443    443        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      444    444        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      455    455        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      470    470        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      483    483        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      487    487        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      490    490        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      511    511        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      512    512        ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      518    518        ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE   1012 AA;  112400 MW;  8D617C4DBF0CB0F7 CRC64;

Query Match                    25.4%;   Score 713.5;  DB 1;  Length 1012;
Best Local Similarity          33.9%;   Pred. No. 4.6e-44;
Matches   185;  Conservative   97;  Mismatches   214;  Indels   49;  Gaps   17;

Qy    2  APKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH   61
         |||  |:  || ||    |   :|  ||    :| | | | : ||      ||  :
```

```
Db   493  APKGKSAAPSKKSK----GLYKEEGVNKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKGG  548
             :  :||    | : ||  ||:    |:.: | |   : :|
Qy    62  EDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLEDA   118
             || |     | :||:||   ||:   ||| | ||
Db   549  KVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVIGSNKLEQMPSKEDA   608
             | :| ::  || :  :|:||| ||   :|: |:
Qy   119  KKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTV   178
             : :|| ||| || :|:     ::  :| |  | :   :|||
Db   609  VEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKK----------LT    654
Qy   179  VKP---CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEAL   235
             |||       |:|| :  :|   ::||:|:||||:||||||||:|:|:||  : |||
Db   655  VKPGTKSKLPKAVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSKRQIQAAYSILSEV   714
             :| |     :            :     |      | ||:|:|  | ::|
Qy   236  EEAMKNPTGDGQSLEELSSCFYTVIPIHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQ  295
             ::| ::   |:|| ||:| |||:||: :| ||   :|| | | ||:| :|: :| ||:
Db   715  QQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAY   773
Qy   296  TLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYR    351
             :|    |:| :  : |:| | | : | :| :  ||| :|  : |:|        |||:|
Db   774  SLLRGGSDDDSSK-----DPIDVVYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYD  828
Qy   352  CPNLRHVKVKNVREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGG  407
             | ||:|:|  |||||| :| || ||:|||||||||: | :|||:|:||||:||      ||
Db   829  L-EVMDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGIISQGLRIAPPEAPVTGY   887
Qy   408  RVGKGIYFASENSKSAGYVTTMHCGGHOVGYMFLGEVALGKEHITIDDPSLKSPPPGFD   467
             :||||||||  | :||||:|| ||| | ||:| :|||||:|:||  :|| |:|||||:|
Db   888  MFGKGIYFADMVSKSANYCHTSQ--GDPIGLILGEVALGNMYELK-HASHISKLPKGKH   944
Qy   468  SVIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLR   526
             || |       |:  | | |    || :|| ||    :| : :|  |:: : :|: ||
Db   945  SVKGLGKTPDPSASITLE----GVEVPLGTGI--PSGVNDTCLLYNEYIVIDIAQVNLK   998
Qy   527  YLLEI   531
             ||| |
Db   999  YLLKL  1003
```

RESULT 3
```
ID   PPO2_MOUSE     STANDARD;      PRT;    559 AA.
AC   O88554; Q99N29;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly(ADP-ribose) synthetase-2) (pADPRT-2)
DE   (mPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Embryo;
RX   MEDLINE=99292755; PubMed-10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129Sv;
RX   MEDLINE=21179160; PubMed-11113988;
RA   Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,
RA   Niedergang C.P.;
RT   "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2
RT   (PARP-2) gene to the gene for RNase P RNA.";
RL   J. Biol. Chem. 276:11092-11099(2001).
RN   [3]
RP   SEQUENCE OF 9-559 FROM N.A.
RC   STRAIN=129/Sv X C57BL/6;
RX   MEDLINE=9926466; PubMed-10338144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC       SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in
CC       testis followed by ovary.
CC   -!- INDUCTION: By high levels of DNA-damaging agents.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ007780; CAA07679.1; -.
DR   EMBL; AF191547; AAK13253.1; -.
DR   EMBL; AF072521; AAC25415.1; ALT_INIT.
     HSSP; P26446; 1A26.
DR   MGD; MGI:1341112; Adprt2.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND      1     65       POTENTIAL.
FT   DOMAIN       66    559       NAD-BINDING (BY SIMILARITY).
FT   DOMAIN        3      9       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                               (POTENTIAL).
FT   DOMAIN       33     39       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                               (POTENTIAL).
FT   CONFLICT     82     82       L -> V (IN REF. 2).
FT   CONFLICT    177    177       V -> I (IN REF. 2).
FT   CONFLICT    486    486       R -> Q (IN REF. 2).
SQ   SEQUENCE   559 AA;  63396 MW;  E0AEDAEE412C1445 CRC64;

Query Match             25.2%;  Score 708;  DB 1;  Length 559;
Best Local Similarity   34.4%;  Pred. No. 5.1e-44;
Matches  200;  Conservative   88;  Mismatches 212;  Indels  82;  Gaps  23;

Qy    1  MAPKRKAS-----VQTEGSKKQRQGTE--EEDS--FRSTAEALRAAP---------ADN   41
            |||:|  |      | :| |||:|| :   :|||    : |: ||||
Db    1  MAPRQRSGSGSRRVLNE-AKKVDNGNKATEDDSPPGKKMRTCQRKGPMAGGKDADRTKDN   59
Qy   42  R--------VIRVDPSCPFSRNPG---IQVHEDYDCTLNQTNIGNNNNKFYIIQLLEE    88
            |          | | | | |      |||:|||||:|||||:|:|||:|||:|:||||
Db   60  RDSVKTLLLKGKAPVDPECAAKLGKAHVICEGDDVYDVMLNQTNLQFNNNKYYLIQLLED  119
Qy   89  GSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFKKKFWEKTKNKWEERDRFVAQ  143
            ::|   : | |||||| ||:|   |  |    ::||||:|| |:| |||:||::|:|: :
Db  120  DAQRNFSVWMRWGRVGKTGQHSL--VTCSGDLNKAKEIFQKKFLDKTKNNWEDRENFEKV  177
Qy  144  PNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLDPATQNLITNIFSKEMFKN  203
            |.|| :::::|:|| || | |||    :     || |   :   :    | ::  :: :
Db  178  PGKYDMLQMDYAASTQD-------ESKTKEETLKPESQLDRVQELLKLICNVQTMEE     229
Qy  204  AMTLMNLDVKKMPLGKLTKQQIARGFEALEAEAMKNPTGD-GQSLEELSSCFYTVIPH    262
            |   : ||::: |||||||  ||:| : || :|   :|   |||:: || | :|| ||
Db  230  MMIEMKYDTKRAPLGKLTVAQIKAGYQSLKKIEDCIR--AGQHGRALVEACNEFYTRIPH  287
```

Qy 263 NFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEEKVEEVPHPLDRDYQLL 322
         :||:|| :|| : :| :|| :|||:| :||:| : +: |||| |:|
Db 288 DFGLSIPPVIRTEKELSDKVKLLEALGDIEIALKLVKS---ERQGLE--HPLDQHYRNL 341

Qy 323 RCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------NLRHVWKVNREGEGDRFQAHSK 375
         :| :||: :|:| :| |:: |:| : :||::|: :|:|
Db 342 HCALRPLDHESNEFKVISQYIQST----HAPTHKDYTMTLLDVFEVEKEGEKEAFR--ED 395

Qy 376 LGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSKSAGYVTTMHC 431
         | ||| |||||: ::|||| |||| : :||:| ||||||||||:|||
Db 396 LPNRMLLWHGSRLSNWVGILSHGLRVAPPEAPITGVMFGKGIYFADMSSKSANYCFASRL 455

Qy 432 GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIELELDGOP 491
         :|: | ||||| : : :| : |||| | || | :|| :|
Db 456 --KNTGLLLLSEVALGQCNELLEANPKAQGLLRGKHSTKGMGKMAPSPAHFI--TLNGS- 510

Qy 492 VVVPQGPPVQCPSFKSSSF--SQSEYLIYKPSQCRLRYLLEI 531
         |:|:| | :::|| :|:|| |:|||| | | :| :||::|
Db 511 -TVPLGPASDTGILNPEGYTLNYNEFIVYSPNQVRMRYLLKI 551

RESULT 4
PPOL_HUMAN
ID   PPOL_HUMAN   STANDARD;   PRT;   1013 AA.
AC   P09874;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT OR PPOL OR PARP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90091744; PubMed=2513174;
RA   Auer B., Nagl U., Herzog H., Schneider R., Schweiger M.;
RT   "Human nuclear NAD+ ADP-ribosyltransferase(polymerizing):
RT   organization of the gene.";
RL   DNA 8:575-580(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88076933; PubMed=3120710;
RA   Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S.,
RA   Suzuki H., Nyunoya H., Miwa M., Sugimura T.;
RT   "Nucleotide sequence of a full-length cDNA for human fibroblast
RT   poly(ADP-ribose) polymerase.";
RL   Biochem. Biophys. Res. Commun. 148:617-622(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88058958; PubMed=2824474;
RA   Kurosaki T., Ushiro H., Mitsuuchi Y., Suzuki S., Matsuda M.,
RA   Matsuda Y., Katunuma N., Kangawa K., Matsuo H., Hirose T.,
RA   Inayama S., Shizuta Y.;
RT   "Primary structure of human poly(ADP-ribose) polymerase as deduced
RT   from cDNA sequence.";
RL   J. Biol. Chem. 262:15990-15997(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88068596; PubMed=2891139;
RA   Cherney B.W., McBride O.W., Chen D., Alkhatib H., Bhatia K.,
RA   Hensley P., Smulson M.E.;
RT   "cDNA sequence, protein structure, and chromosomal location of the
RT   human gene for poly(ADP-ribose) polymerase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:8370-8374(1987).
RN   [5]
RP   SEQUENCE OF 440-1013 FROM N.A.
RX   MEDLINE=87298455; PubMed=3113420;
RA   Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,

RA   Miwa M.;
RT   "Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and
RT   expression of its gene during HL-60 cell differentiation.";
RL   Biochem. Biophys. Res. Commun. 146:403-409(1987).
RN   [6]
RP   ERRATUM.
RA   Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA   Miwa M.;
RL   Biochem. Biophys. Res. Commun. 148:1549-1550(1987).
RN   [7]
RP   SEQUENCE OF 1-94 FROM N.A.
RX   MEDLINE=91099327; PubMed=2125269;
RA   Yokoyama Y., Kawamoto T., Mitsuuchi Y., Kurosaki T., Toda K.,
RA   Ushiro H., Terashima M., Sumimoto H., Kuribayashi I., Yamamoto Y.,
RA   Maeda T., Ikeda H., Sagara Y., Shizuta Y.;
RT   "Human poly(ADP-ribose) polymerase gene. Cloning of the promoter
RT   region.";
RL   Eur. J. Biochem. 194:521-526(1990).
RN   [8]
RP   SEQUENCE OF 1-39 FROM N.A.
RX   MEDLINE=90211250; PubMed=2108670;
RA   Ogura T., Nyunoya H., Takahashi-Masutani M., Miwa M., Sugimura T.,
RA   Esumi H.;
RT   "Characterization of a putative promoter region of the human
RT   poly(ADP-ribose) polymerase gene: structural similarity to that of
RT   the DNA polymerase beta gene.";
RL   Biochem. Biophys. Res. Commun. 167:701-710(1990).
RN   [9]
RP   SEQUENCE OF 1-39 FROM N.A.
RA   Herzog H., Schneider R., Hirsch-Kauffmann M., Schnitzer D.,
RA   Schweiger M.;
RL   Submitted (JUL-1991) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=90222155; PubMed=2109322;
RA   Gradwohl G., Menissier de Murcia J., Molinete M., Simonin F.,
RA   Koken M.H.M., Hoeijmakers J.H.J., de Murcia G.M.;
RT   "The second zinc-finger domain of poly(ADP-ribose) polymerase
RT   determines specificity for single-stranded breaks in DNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:2990-2994(1990).
RN   [11]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=91072398; PubMed=2123876;
RA   Ikelima M., Noguchi S., Yamashita R., Ogura T., Sugimura T.,
RA   Gill D.M., Miwa M.;
RT   "The zinc fingers of human poly(ADP-ribose) polymerase are
RT   differentially required for the recognition of DNA breaks and nicks
RT   and the consequent enzyme activation. Other structures recognize
RT   intact DNA.";
RL   J. Biol. Chem. 265:21907-21913(1990).
RN   [12]
RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=91035460; PubMed=2121735;
RA   Simonin F., Menissier de Murcia J., Poch O., Muller S., Gradwohl G.,
RA   Molinete M., Penning C., Keith G., de Murcia G.M.;
RT   "Expression and site-directed mutagenesis of the catalytic domain of
RT   human poly(ADP-ribose)polymerase in Escherichia coli. Lysine 893 is
RT   critical for activity.";
RL   J. Biol. Chem. 265:19249-19256(1990).
RN   [13]
RP   NUCLEAR LOCALIZATION SIGNAL.
RX   MEDLINE=92371433; PubMed=1505517;
RA   Schreiber V., Molinete M., Boeuf H., de Murcia G.M.,
RA   Menissier de Murcia J.;
RT   "The human poly(ADP-ribose) polymerase nuclear localization signal is
RT   a bipartite element functionally separate from DNA binding and
RT   catalytic activity.";
RL   EMBO J. 11:3263-3269(1992).
RN   [14]
RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=97461532; PubMed=9315851;
RA   Rolli V., O'Farrell M., Menissier-de Murcia J., de Murcia G.M.;
RT   "Random mutagenesis of the poly(ADP-ribose) polymerase catalytic

RT   domain reveals amino acids involved in polymer branching.";
RL   Biochemistry 36:12147-12154(1997).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor -
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
DR   EMBL; X16674; CAA34663.1; -.
DR   EMBL; M18112; AAA60137.1; -.
DR   EMBL; J03473; AAB59447.1; -.
DR   EMBL; M17081; AAA51599.1; ALT_SEQ.
DR   EMBL; M32721; AAA60155.1; -.
DR   EMBL; M29786; AAA51663.1; -.
DR   EMBL; M29545; AAA51663.1; JOINED.
DR   EMBL; M29766; AAA51663.1; JOINED.
DR   EMBL; M29767; AAA51663.1; JOINED.
DR   EMBL; M29768; AAA51663.1; JOINED.
DR   EMBL; M29769; AAA51663.1; JOINED.
DR   EMBL; M29770; AAA51663.1; JOINED.
DR   EMBL; M29771; AAA51663.1; JOINED.
DR   EMBL; M29772; AAA51663.1; JOINED.
DR   EMBL; M29773; AAA51663.1; JOINED.
DR   EMBL; M29774; AAA51663.1; JOINED.
DR   EMBL; M29775; AAA51663.1; JOINED.
DR   EMBL; M29777; AAA51663.1; JOINED.
DR   EMBL; M29778; AAA51663.1; JOINED.
DR   EMBL; M29779; AAA51663.1; JOINED.
DR   EMBL; M29780; AAA51663.1; JOINED.
DR   EMBL; M29781; AAA51663.1; JOINED.
DR   EMBL; M29783; AAA51663.1; JOINED.
DR   EMBL; M29784; AAA51663.1; JOINED.
DR   EMBL; M29785; AAA51663.1; JOINED.
DR   EMBL; M29544; AAA51663.1; JOINED.
DR   EMBL; M29782; AAA51663.1; JOINED.
DR   EMBL; X56140; CAA39606.1; -.
DR   EMBL; X56141; CAA39606.1; JOINED.
DR   EMBL; M60436; AAA60000.1; -.
DR   PIR; A26901; A26901.
DR   PIR; A28498; A28498.
DR   PIR; A29725; A29725.
DR   PIR; A35635; A35635.
DR   PIR; A33321; A33321.
DR   PIR; B33321; B33321.
DR   PIR; A39976; A39976.
DR   PIR; S14010; S14010.
DR   HSSP; P26446; 1A2E.
DR   Aarhus/Ghent-2DPAGE; 1620; NEPHGE.
DR   MIM; 173870; -.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.

Query Match   25.1%; Score 705; DB 1; Length 1013;
Best Local Similarity   33.8%; Pred. No. 1.9e-43;
Matches 183; Conservative 101; Mismatches 218; Indels 40; Gaps 17;

```
Qy    1  MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV   60
Db  492  VAPRGKSGA---ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG  548

Qy   61  HEDYDCTLNQTNIGNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED   117
Db  549  GKVFSATGLGVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKED  608

Qy  118  AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT  177
Db  609  AIEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL---TVNPGTKSK  663

Qy  178  VVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE  237
Db  664  LPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQ  717

Qy  238  AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTL   297
Db  718  AVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPLLNNADSVQAKVEMLDNLLDIEVAYSL   776

Qy  298  QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYRCP   353
Db  777  LRGGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATTHNAYDL-  830

Qy  354  NLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRV   409
Db  831  EVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMF  890

Qy  410  GKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSV  469
Db  891  GKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSV  947

Qy  470  IARGQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLLYKESQCRLRYLL   529
Db  948  KGLGKTTPDPSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEYIVYDIAQVNLKYLL 1002

Qy  530  EI  531
Db 1003  KL 1004
```

RESULT   5
PPOL_BOVIN
ID   PPOL_BOVIN   STANDARD;   PRT;   1015 AA.
AC   P18493; Q9TS00;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90382673; PubMed=2119324;
RA   Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;
RT   "Cloning of a full-length cDNA encoding bovine thymus
RT   poly(ADP-ribose) synthetase: evolutionarily conserved segments and
RT   their potential functions.";
RL   Gene 90:249-254(1990).

```
RN   [2]
RP   SEQUENCE OF 647-714 AND 838-903 FROM N.A.
RX   MEDLINE=88151954; PubMed=2450019;
RA   Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,
RA   Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;
RT   "Depression in gene expression for poly(ADP-ribose) synthetase during
RT   the interferon-gamma-induced activation process of murine macrophage
RT   tumor cells.";
RL   Eur. J. Biochem. 171:571-575(1988).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; D90073; BAA14114.1; -.
DR   EMBL; X06986; CAA30046.1; -.
DR   EMBL; X06987; CAA30047.1; -.
DR   PIR; JS0428; JS0428.
DR   PIR; S00328; S00328.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   DNA_BIND      1    374
FT   DOMAIN      375    525       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      386    462       BRCT.
FT   DOMAIN      526   1015       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     127    164       PARP-TYPE.
FT   DOMAIN      209    211       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     408    408       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     414    414       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     446    446       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     449    449       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     457    457       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     472    472       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     489    489       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     492    492       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     514    514       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     515    515       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     521    521       ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE   1015 AA;  113355 MW;  0A5FE9D9F04F5B04 CRC64;

Query Match             25.0%;  Score 702;  DB 1;  Length 1015;
Best Local Similarity   34.2%;  Pred No. 3.2e-43;
Matches  187;  Conservative   95;  Mismatches  215;  Indels   50;  Gaps   20;

Qy      1  MAPKRKASVQTEGSKKQRQGTEEDSFRSTAE--ALRAAPADNRVIRVDPSCPFSRNPG  57
              : || ||  |||  : ||   ||: :||  :                 ||
Db    494  VGPKGKSGAAP--SKKSKGPVKEEGTNKSEKRMKLTLKGGAA------VDPDSGLEHNAH  545

Qy     58  I--QVHEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHF  112
              :  || :|   | :|:|||| : | | |||  |  ||||| | ::
Db    546  VLEKGGKVFSATLGLVDIVKGTNSYKLQLLEDDKESRYWIFRSWGRVGTVIGSNKLEQM  605

Qy    113  TCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDS  172
            |||| | ::  | |: |:| |: :  |: :| |  || | |: ||| |  :|
Db    606  PSKEDAIEHFMKLYEKTGNAWHSKN-FTKHPKKFYPLEID-YGQDEEAVKKL---TVNP  660

Qy    173  GPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL  232
            |  |||: :|| |  |: |||:|:: |:|
Db    661  GTKSKLPKP------VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSIL  714

Qy    233  EALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIE  292
            ::|:  :|  | |||:||:| ::|::| ::|| : |||| ||  :|| :|| |||:
Db    715  SEVQQALSQGSSDSHIL-DLSNRFYTLIPHDFGMKKPPLLNNANSVQAKVEMLDNLLDIE  773

Qy    293  LAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGN  348
            :|: ||  |  |:  ||   :| |  ||: |: :| :
Db    774  VAYSLLRGGSDDSSK-----DPIDVNYEKLKTDIKVVDKDSEEAEIIRKYVKNTHATTHN  828

Qy    349  SYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----  404
            :|  |||:||| : ::  ||: |:| ||||||||:
Db    829  AYDL-EVVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPV  887

Qy    405  SGGRVGKGIYFASENSKSAGYVTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPP  464
            || |||||||| |:| ||    | :  || |   :|: :|||:|| :|| :|| :
Db    888  TGYMFGKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEAALGNMYELK-HARHISKLPK  944

Qy    465  GFDSVIARGQTEPPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCR  524
            | || |  |  | ||  :  :| | ||:: :||:| : :|: |: ::|
Db    945  GKHSVKGLGKTTPDPSASI--TVDG--VEVPLGTGIS-SGVNDTCLLYNEYIVYDIAQVH  999

Qy    525  LRYLLEI  531
            |:||:::
Db   1000  LKYLLKL  1006

RESULT   6
PPO2_HUMAN
ID   PPO2_HUMAN     STANDARD;      PRT;   583 AA.
AC   Q9UGN5; Q9Y6C8; Q9NDV2; Q9UMR4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (PADPRT-2)
DE   (hPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal brain;
```

```
FT   VARSPLIC      68      80       MISSING (IN ISOFORM 2).
FT   CONFLICT     447     447       P -> H (IN REF. 2).
FT   CONFLICT     481     481       N -> H (IN REF. 4).
SQ   SEQUENCE    583 AA;    66205 MW;   5B7AE8AE531836AF CRC64;

Query Match           24.8%;    Score 695.5; DB 1;   Length 583;
Best Local Similarity  33.1%;   Pred. No. 3.7e-43;
Matches  185; Conservative  97; Mismatches 200; Indels  77; Gaps  19;

Qy     13 GSKKQRQGTEEED--------SFRSTAEALRA------APADNRVIRVDPSCPFSRNPG-    57
                                                               .
Db     54 GGKANKDRTEDKQDGMPGRSWASKRVSESVKALLLKGKAP------VDPECTAKVGKAH   106

Oy     58 --IQVHEDYDCTLNQTNIGNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFT   113
              :: || ||||||  :: | |||||:||:|| ||||||::||
Db    107 VYCEGNDVYDVMLNQTNLQFNNNKYYIIQLLEDDAQRNFSVMWRWGRVGKMGQHSL--VA   164

Qy    114 C---LEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEV------QGEAESQEAV   162
             | :: |||| |||||  :: |||  :::     || ::    |:||
Db    165 CSGNLNKAKEIFQKKFLDKTKNWEDREKFEKVPGKYDMLQMDYATNTQDEETKKEESL   224

Qy    163 VKALSPQVDSGPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTK   222
                                                        ||||| :|| |||:|||
Db    225 KSPLKPE---------------SQLDLRVQELIKLLICNVQAMEEMMEMKYNTKKAPLGKLTV   272

Qy    223 QQIARGFEALEALEAEAMKNPTGD-GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAK   281
              || :|:::: :|::.:. :: || :|||||:::|||:||:||:||| :|: ||||||
Db    273 AQIKAGYQSLKKIEDCIR--AGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEK   330

Qy    282 KDMLLVLADIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQT   341
               :| ||||:||:|  ::: :::    |||||| :| ||:|:|||:||||
Db    331 IQLLEALGDIEIAIKL-------VKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQ   384

Qy    342 YLKQT---GNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSG   398
               :||::  | |||| |:|:|:. ||::::: .| :|.|:    | |||:   :||
Db    385 YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHG   442

Qy    399 LRIMPH---SGGRVGKGIYFASENSKSAGYYTTMHCGGHQVGYMFLGEVALGKEHHITI   454
               ||| :    |||||||||:     ::  || ||:::|:: |:   ||:|||| |||:
Db    443 LRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRL--KNTGLLLLSEVALGQCNELLE   500

Qy    455 DDPSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSF--SQ   512
              : :: ||  || |:|:|||:|:  ||| ::||| :|| : || | ||
Db    501 ANPKAEGLLQGKHSTKGLGKMAPSSAHFV--TLNGS--TVPLGPASDTGILNPDGYTLNY   556

Qy    513 SEYLIYKESQCRLRYLLEI   531
           :||::|:: :|::|||:::
Db    557 NEYIVYNPNQVRMRYLLKV   575
```

RESULT 7
PPOL_RAT

ID   PPOL_RAT   STANDARD;   PRT;   1013 AA.
AC   P27008; O35937;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Monocytes;
RX   MEDLINE=98045546; PubMed=9385436;
RA   Beneke S., Meyer R., Buerkle A.;
RT   "Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly
RT   (ADP-ribose) polymerase.";
RL   Biochem. Mol. Biol. Int. 43:755-761(1997).

RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoeger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE OF 2-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [3]
RP   SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fibroblast;
RX   MEDLINE=99268466; PubMed=10338144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA   Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA   Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA   Masuho Y., Kanehori K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC   SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (BY
CC   similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC   produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN
CC   THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO
CC   DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;
CC   LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND
CC   THYMUS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AJ236912; CAB65088.1; -.
DR   EMBL; AF085734; AAD29857.1; ALT_INIT.
DR   EMBL; AJ236876; CAB41505.2; ALT_INIT.
DR   EMBL; AK001980; BAA92017.1; ALT_TERM.
HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Alternative splicing.
FT   DNA_BIND   1   88   POTENTIAL.
FT   DOMAIN   86   583   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   4   7   NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT   (POTENTIAL).
FT   DOMAIN   35   40   NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT   (POTENTIAL).

```
FT   DOMAIN       206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN       220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES        1      1       BLOCKED (BY SIMILARITY).
FT   MOD_RES      407    407       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      413    413       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      435    435       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      437    437       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      456    456       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      484    484       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      491    491       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      513    513       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      519    519       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT     638    638       Y -> H (IN REF. 4).
FT   CONFLICT     641    641       E -> A (IN REF. 4).
FT   CONFLICT     752    752       N -> D (IN REF. 4).
SQ   SEQUENCE   1013 AA;  112529 MW;  AA566F2B29BE97C0 CRC64;

Query Match        24.7%;
Best Local Similarity 33.5%;            Score 694.5;  DB 1;  Length 1013;
Matches  182;  Conservative 101;  Mismatches 216;  Indels 45;  Gaps 18;
                                        Pred. No. 1.1e-42;

Qy      1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALARAAPADNRVIRVDPSCPFSRNPGIQV   60
            ||:: || ::  :   ||  :  : :        :: :|      :  ::   : |
Db    493 VVPKGKSAAPSKKSK----GAVKEEGVNKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG  548

Qy     61 HEDYDCTLNQTNIGNNNNKFYIIQLLE-EGSRFFCWNRWGRVGEV-GQSKMNHFTCLED   117
            : :| : || |  :|||| | :||||| ||  | | ::|||:| || :| |: :| :|||
Db    549 GKVFSATLGLVDIVKGTNSYYKLQLLESDKESRYWIFRSWGRVGTVIGSNKLEQMPSKED  608

Qy    118 AKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK-ALSPQVDSGPVR   176
            | :|  | | ||  |||  | | :::  |  ||  | :  |||: |  | ::   | |:
Db    609 AVEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKLAVKPGTKS----  662

Qy    177 TVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQTARGFEALEALE   236
            :  :  :||  ::| :    :::|  | : ::  : |:|||||| : :|:  |:
Db    663 ------KLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSEVQ  716

Qy    237 EAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQT   296
            :|:   ::: | | ::  :|||:||| ||  :|| :| |:||||::::::||: |||::
Db    717 QAVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNTDSVQAKVEMLDNLLDIEVAYS  775

Qy    297 LQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRC   352
            |  |    ::  : ::| || ||   : | ||:  | ||  | : |||     || : |
Db    776 LLRGGSDDSSK------DPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDL  830

Qy    353 PNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGR   408
            :  : :|: ||||| :|: |: :|  |||||||: ::: | || |||| ||     |  |
Db    831 -EVIDFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYM   889

Qy    409 VGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDS   468
            .|||||||  :|||| | :|:|    |:| |:||||||| :  : :    : ||| |:|:
Db    890 FGKGIYFADMVSKSANYCHTSQ-GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHS   946

Qy    469 VIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRLRY   527
            | :|  | ||||:|  : ||| | ||  |  | :    : :   : :|::| :::  |:|
Db    947 VKGLGKTAPDPSASI--TLDG--VEVPLGTGI--PSGVNDTCLLYNEYIVYDIAQVNLKY  1000

Qy    528 LLEI  531
            ||::
Db   1001 LLKL 1004

RESULT 8
PPOL_CHICK
ID   PPOL_CHICK     STANDARD;      PRT;   1011 AA.
AC   P26446;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
```

```
RN   [2]
RP   REVISION TO 811.
RA   Beneke S., Meyer R., Buerkle A.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-11 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=9229013; PubMed=1601134;
RA   Potvin F., Thibodeau J., Kirkland J.B., Dandenault B.,
RA   Duchaine C., Poirier G.G.;
RT   "Structural analysis of the putative regulatory region of the rat
RT   gene encoding poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 302:269-273(1992).
RN   [4]
RP   SEQUENCE OF 514-1013 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=90027702; PubMed=2508731;
RA   Thibodeau J., Gradwohl G., Dumas C., Clairoux-Moreau S., Brunet G.;
RT   "Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase
RT   catalytic domain and analysis of mRNA levels during the cell cycle.";
RL   Biochem. Cell Biol. 67:653-660(1989).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U94340; AAC53544.1; -.
DR   EMBL; X65496; CAA46477.1; -.
DR   EMBL; X65497; CAA46478.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS00347; PARP_ZN_FINGER_2; 2.
DR   PROSITE; PS50064; NAD; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET       0      0
FT   DNA_BIND       1    372       BY SIMILARITY.
FT   DOMAIN       385    461       BRCT.
FT   DOMAIN       373    523       AUTOMODIFICATION DOMAIN.
FT   DOMAIN       524   1013       NAD-BINDING.
FT   ZN_FING       20     55       PARP-TYPE.
FT   ZN_FING      124    161       PARP-TYPE.
```

```
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=91340148; PubMed=1840535;
RA   Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;
RT   "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid
RT   sequence and comparison with mammalian enzyme sequences.";
RL   Gene 102:157-164(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.
RX   MEDLINE=96353841; PubMed=8755499;
RA   Ruf A., Mennissier de Murcia J., de Murcia G.M., Schulz G.E.;
RT   "Structure of the catalytic fragment of poly(AD-ribose) polymerase
RT   from chicken.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO
RP   895.
RX   MEDLINE=98191351; PubMed=9521710;
RA   Ruf A., de Murcia G.M., Schulz G.E.;
RT   "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived
RT   from crystal structures and homology modeling.";
RL   Biochemistry 37:3893-3900(1998).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.
RX   MEDLINE=98239716; PubMed=9571033;
RA   Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;
RT   "The mechanism of the elongation and branching reaction of poly(ADP-
RT   ribose) polymerase as derived from crystal structures and
RT   mutagenesis.";
RL   J. Mol. Biol. 278:57-65(1998).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```

```
DR   EMBL; X52690; CAA36917.1; -.
DR   PIR; JH0581; JH0581.
DR   PDB; 2PAW; 27-MAY-98.
DR   PDB; 1PAX; 15-MAY-97.
DR   PDB; 2PAX; 27-MAY-98.
DR   PDB; 3PAX; 27-MAY-98.
DR   PDB; 4PAX; 27-MAY-98.
DR   PDB; 1A26; 27-MAY-98.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP.reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; 3D-structure.
FT   DNA_BIND      1    370
FT   DOMAIN      371    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      382    458       BRCT.
FT   DOMAIN      523   1011       NAD-BINDING.
FT   ZN_FING      21     56       PARP-TYPE.
FT   ZN_FING     125    162       PARP-TYPE.
FT   DOMAIN      207    209       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     403    403       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     404    404       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     410    410       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     411    411       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     432    432       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     441    441       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     442    442       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     453    453       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     454    454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     468    468       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     481    481       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     509    509       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     510    510       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     517    517       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    895    895       A -> R (IN REF. 1).
SQ   SEQUENCE   1011 AA;  113520 MW;  261AED9383139144 CRC64;

Query Match             24.6%;  Score 691.5;  DB 1;  Length 1011;
Best Local Similarity   33.9%;  Pred. No. 1.8e-42;
Matches  185;  Conservative  97;  Mismatches 206;  Indels  57;  Gaps  20;

Qy      5 RKASVQTEGSKKQRQG-TEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED 63
Db    497 KPANMKSAGKVKEEQGPSKSEKKMLTVKGGAAVDPDSGL-EDSAHVFEKGGKI------ 549

Qy     64 YDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKK 120
Db    550 FSATIGLVDIVKGTNSYYKIQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAVE 609

Qy    121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVK 180
Db    610 HFLNLYEEKTGNSWHSKN-FTKYPKKFYPLEID-YGQDEEAVRKL---TVSAGTKSKLAK 664

Qy    181 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 240
Db    665 P------IQDLIKMIFDVESMKKAWVEFEIDLQKRMPLGKLSKRQIQSAYSILNEVQQAVS 718

Qy    241 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAA 300
Db    719 DGGSESQIL-DLSNRFYTLLPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRG 777

Qy    301 PGEEEKVEEVPHPLDRDYQLLRCOLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLR 356
Db    778 GNEDGDK-----DPIDINYEKLRTDIKVVDKDSEEAKIIKQYVKNTHAATHNAYDL-KVV 831
```

```
SQ  SEQUENCE    637 AA;   72175 MW;   527A8F464605D127 CRC64;

Query Match                24.5%;  Score 690;  DB 1;  Length 637;
Best Local Similarity      34.7%;  Pred. No. 1.2e-42;
Matches 190;  Conservative  90;  Mismatches 219;  Indels  48;  Gaps  18;

Qy      6 KASVQTEGSKKQRQGTEEDSFRS--TAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED  63
            |:|:||     | ||:::| |   ::|||||       : ||||||   :|||
Db    112 KSSNGTDEAEDDNNGFEEEKKEKIVTATKKGAAVLDQWI---PDEIKSQYHVLQRGDD 167

Qy     64 -YDCTLNQTNIGNNNKFYIIQLLEEGSR--FFCWNRRGRVGEVGQSKMN-HFTCLEDAK 119
           |   ||||||  ::||:|        |      ||||||||  |:  |||||
Db    168 VYDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVYTRWGRVGVKGQSKLDGPYDSWDRAI 227

Qy    120 KDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVV 179
          ::|:   |||||   | | ||| |  ::  || |     ::  ||     | | ::
Db    228 EIFTNKFNDKTKNYWSDRKEFIPHPKSVTWLEMDYGKEENDSPVNNDIPSSSS-----EV 282

Qy    180 KP--CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
           ||    ::|  :| |:::| :|   ||||:|| ||| || :::| |:::::|  ::||
Db    283 KPEQSKLDTRVAKFISLICNVSMAQHMEIGYNANKLPLGKISKSTISKGYEVLKRISE 342

Qy    238 AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPP--INSPDVLQAKKDMLLVLADIELA- 294
           :                     ||||||:||  |  :|:|||:|||   :::|||||
Db    343 VIDR--YDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELAT 400

Qy    295 QTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYR 351
            :  || :|  :      |  |: |:| ||     : |:
Db    401 KLLSVDPGLQDD------PLYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGY 453

Qy    352 CPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGG 407
          ||| ||| ||| | | | | ||| |:||||||||:|| ||| |||||||    ||
Db    454 TVEIAQLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGY 513

Qy    408 RVGKGIYFASENSKSAGYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPP 464
           | ||||||:|||:|||| |||:|   |||| || |||||| | | |  |||||| |||
Db    514 MFGKGVFADMFSKSANY-----CYANTGANDGVLLCEVALGDMNELLYSDYNADNLPP 568

Qy    465 GFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCR 524
           | |:|   |   ||  ||:|:|||||||| :: ||||  ||:::|| |:||||||||:|
Db    569 GKLSTKGVGKTAPNPSEAQTLE-DG--VVVPLGKPVE-RSCSKGMLLYNEYIVYNVEQIK 624

Qy    525 LRYLLEI 531
          ||:||:|
Db    625 MRVYIQV 631
```

```
RESULT 9
PPOL_ARATH     STANDARD;     PRT;    637 AA.
ID   PPOL_ARATH
AC   Q11207;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   APP.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95269779; PubMed=7750552;
RA   Lepiniec L., Babiychuk E., Kushnir S., van Montagu M., Inze D.;
RT   "Characterization of an Arabidopsis thaliana cDNA homologue to animal
RT   poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 364:103-108(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC       nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z48243; CAA88288.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND      1    140       POTENTIAL.
FT   DOMAIN      140    637       NAD-BINDING (BY SIMILARITY).
FT   DOMAIN       41     62       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
```

```
Qy    357 HVKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKG 412
          :::: |||  ||   : ||:||||| : | ||||||||  |||||      |||:|||
Db    832 EIFRIEREGESQRYKPFKQLHNRQLLWHGSRTNFAGILSQGLRIAPPEAPVTGYMFGKG 891

Qy    413 IYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEH-----HITIDDPSLKSPPPGF 466
          ||||     | : |||        |||||||||||||          |||
Db    892 IYFADMVSKSANYCHTSQ--ADPIGLILLGEVALGNMYELKNASHIT------KLPKGK 942

Qy    467 DSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLR 526
           |  ||| | || | :: ||       |   |||  ||:||:| :| :  |||
Db    943 HSVKGLGKTAPDPT--ATTTLDG--VEVPLGNGIS-TGINDTCLLYNEYIYVDVAQVNLK 997

Qy    527 LRYLLEI 531
          |||::
Db    998 YLLKL 1002
```

```
RESULT 10
PPOL_XENLA
ID   PPOL_XENLA     STANDARD;     PRT;    998 AA.
AC   P31669;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Saulier-Le Drean B.M.;
RL   Thesis (1992), University of Rennes, France.
RN   [2]
RP   SEQUENCE OF 742-876 FROM N.A.
RX   MEDLINE=93277538; PubMed=8503897;
RA   Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,
RA   Miwa M.;
RT   "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)
```

```
RT    polymerase from Xenopus laevis and cherry salmon using heterologous
RT    oligonucleotide consensus sequences.";
RL    Biochem. Biophys. Res. Commun. 193:119-125(1993).
CC    -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC        PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC        ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC        CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC        TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC        EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC    -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC        nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
CC    -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC    -!- SUBCELLULAR LOCATION: Nuclear.
CC    -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,
CC        AND BRAIN. LOW IN LIVER.
CC    -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC        AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC        FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC        THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC        AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC    -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC    -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC    --------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------
DR    EMBL; Z12139; CAA78126.1; -.
DR    EMBL; D13810; BAA02966.1; -.
DR    PIR; S31735; S31735.
DR    HSSP; P26446; 1A26.
DR    InterPro; IPR001357; BRCT.
DR    InterPro; IPR001290; PARP.
DR    InterPro; IPR004102; PARP_reg.
DR    InterPro; IPR001510; Znf-PARP.
DR    Pfam; PF00533; BRCT; 1.
DR    Pfam; PF00644; PARP; 1.
DR    Pfam; PF02877; PARP_reg; 1.
DR    Pfam; PF00645; zf-PARP; 2.
DR    ProDom; PD004675; Znf-PARP; 2.
DR    SMART; SM00292; BRCT; 1.
DR    PROSITE; PS50172; BRCT; 1.
DR    PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR    PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW    Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW    ADP-ribosylation; Zinc-finger; Zinc.
FT    NON_TER       1       1
FT    DNA_BIND     <1     356
FT    DOMAIN      357     507       AUTOMODIFICATION DOMAIN.
FT    DOMAIN      369     445       BRCT.
FT    DOMAIN      508     998       NAD-BINDING.
FT    ZN_FING       8      43       PARP-TYPE.
FT    ZN_FING     111     148       PARP-TYPE.
FT    DOMAIN      193     195       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT    DOMAIN      207     212       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT    MOD_RES     391     391       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     397     397       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     419     419       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     428     428       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     429     429       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     445     445       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     447     447       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     454     454       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     467     467       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     471     471       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     477     477       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     495     495       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     496     496       ADP-RIBOSYL[N] (POTENTIAL).
FT    MOD_RES     503     503       ADP-RIBOSYL[N] (POTENTIAL).
```

```
FT    CONFLICT    746     746       Q -> E (IN REF. 2).
SQ    SEQUENCE    998 AA;  111126 MW;  F5A25E4A3366BAE7 CRC64;
```

```
Query Match                                24.4%;    Score 686.5;    DB 1;    Length 998;
Best Local Similarity            34.1%;     Pred. No. 4.2e-42;
Matches   182;   Conservative   95;   Mismatches 212;   Indels   45;   Gaps   19;

Qy     10  QTEGSKKQRQGT-EEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHEDYDCTL  68
               :    : :  : ::  : :  :  : :     : :   :   :::  :  : ::: : 
Db    488  KSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDS---ELEDSCHVLETGG---KIFSATL  540

Qy     69  NQTNIGNNNNKFYIIQLE--EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKKDFKKK  125
               :  :     :::  : :   : :: :::     : :   : :  :  ::   ::  :: 
Db    541  GLVDITRGTNSYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKEDAIEHFLNL  600

Qy    126  FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLD  185
                : ::  :  :  :       ::   :   : ::::: :   ::: ::  :       : 
Db    601  YQDKTGNAWHS-PNFTKYPKKFYPLEI--DYGQEEDVVKKLS--VGAGTKSKLAKP----  651

Qy    186  PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD  245
               :  : ::::   :  ::   : ::::  :  ::  ::  ::: ::  : :  :  :  : 
Db    652  --VQELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAVSESLSE  709

Qy    246  GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE  305
               :   ::  :   :     :   ::       :    :        :  ::   :  :: 
Db    710  ARLL-DLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDG  768

Qy    306  EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLK----QTGNSYRCPNLRHVWKV  361
               ::  :   :  :    : :    :    :  ::   :      ::::   : :::: : 
Db    769  EK----DPIDVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNAYDLEVL-EIFKI  822

Qy    362  NREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFAS  417
               :::: :   : :  ::  :::::: :  :  :: ::::: :     ::::::::: :: 
Db    823  DREGEYQRYKPFKQLHNRQLLWHGSRTTNFAGLLSQGLRIAPPEAPVTGYMFGKGIYFAD  882

Qy    418  ENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHITIDDPSLKSPPPGFDSVIARGQTEP  477
               ::::::  :  : ::  : :  ::  ::::  ::  :: ::::  :: :: ::: :: 
Db    883  MVSKSANYCHAM--PGSPIGLILLGEVALGNMHELKAASQITKL-PKGKHSVKGLGRTAP  939

Qy    478  DPAQDIELELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEI  531
               :  :  ::   :  : :::: : ::   ::     :::::::: :::: ::::: 
Db    940  DPSATV--QLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLKYLLKL  988
```

```
RESULT 11
PPOL_MOUSE
ID    PPOL_MOUSE      STANDARD;      PRT;   1012 AA.
AC    P11103; Q9JLX4; Q9QVQ3;
DT    01-JUL-1989 (Rel. 11, Created)
DT    01-OCT-1989 (Rel. 12, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
DE    Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE    ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1) (msPARP).
GN    ADPRT OR ADPRT1 OR ADPRP.
OS    Mus musculus (Mouse).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX    NCBI_TaxID=10090;
RN    [1]
RP    SEQUENCE FROM N.A. (ISOFORM 1).
RC    STRAIN=BXSB;
RX    MEDLINE=89263780; PubMed=2498841;
RA    Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;
RT    "Sequence and organization of the mouse poly (ADP-ribose) polymerase
RT    gene.";
RL    Nucleic Acids Res. 17:3387-3401(1989).
RN    [2]
RP    SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC    STRAIN=129/Sv X C57BL/6; TISSUE=Fibroblast;
RX    MEDLINE=20270268; PubMed=10809783;
RA    Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;
RT    "Characterization of sPARP-1. An alternative product of PARP-1 gene
```

```
RT   with poly(ADP-ribose) polymerase activity independent of DNA strand
RT   breaks.";
RL   J. Biol. Chem. 275:15504-15511(2000).
RN   [3]
RP   KNOCK-OUT.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jaeger S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC       short form/sPARP-1; may be produced by alternative initiation.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; X14206; CAA32421.1; -.
DR   EMBL; AF126717; AAF61293.1; ALT_INIT.
DR   PIR; S04200; S04200.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1340806; Adprt1.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS500064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative initiation.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1   1012       POLY [ADP-RIBOSE] POLYMERASE-1, LONG
FT                                ISOFORM.
FT   CHAIN       521   1012       POLY [ADP-RIBOSE] POLYMERASE-1, SHORT
FT                                ISOFORM.
FT   INIT_MET    521    521       FOR SHORT ISOFORM.
FT   DNA_BIND      1    371
FT   DOMAIN      372    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      384    460       BRCT.
FT   DOMAIN      523   1012       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       PARP-TYPE.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     406    406       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     412    412       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     443    443       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     447    447       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     455    455       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     483    483       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     487    487       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     490    490       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     511    511       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     518    518       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     590    590       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    607    607       L -> V (IN REF. 2).
FT   CONFLICT    611    611       E -> D (IN REF. 2).
FT   CONFLICT    628    628       Q -> H (IN REF. 2).
FT   CONFLICT    678    678       N -> D (IN REF. 2).
FT   CONFLICT    702    702       D -> E (IN REF. 3).
FT   CONFLICT    716    716       R -> F (IN REF. 3).
FT   CONFLICT    757    757       Q -> L (IN REF. 2).
FT   CONFLICT    856    856       R -> F (IN REF. 3).
FT   CONFLICT    981    981       A -> C (IN REF. 2).
SQ   SEQUENCE   1012 AA;  112968 MW;  4354C3E5F01B9439 CRC64;
```

```
Query Match             24.3%;  Score 683.5; DB 1;  Length 1012;
Best Local Similarity   32.4%;  Pred. No. 7e-42;
Matches  177;  Conservative  105;  Mismatches  215;  Indels   49;  Gaps   17;

Qy     1  MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV   60
             :||:||: ::||  :  : ::|:  |  ::|:   :|  :|      |  |    |  |  :
Db   492  VAPRGKSAAPSKKSK---GCFKEEGVNKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG  547

Qy    61  HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED  117
          : || |    : || ||:|:  ||||  |  :|:||:|:|:|:||: |:||:| ||:||:
Db   548  GKVFSATLGLVDIVKGTNSYKLQLLEDDKESRYWIFRSWGRLGTVIGSNKLEQMPSKEE  607

Qy   118  AKKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESOEAVVKALSPQVDSGPVRT  177
          |  ||:|:|| | || :  ::: | |  :  ||  |    |  ::|| ||  ||| |
Db   608  AVEQFMKLYEETGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKK-----------L  653

Qy   178  VVKP---CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA  234
          |||    |||:|:| ||| |  :  ::  |: : ::|||||||||||::||  |  | | |
Db   654  TVKPGTKSKLPKPVQELVGMIFDVDSMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSE  713

Qy   235  LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELA  294
          : :  |  |  |  ::|  |:||  :|  ||: |:|:||:  |  |:|:|::|: |::|:
Db   714  VQQPVSGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVA  772

Qy   295  QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSY  350
          ::|  |     :   | :      : || |: :|  |: ::| |   ||      :  ||
Db   773  YSLLRGGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAY  827

Qy   351  RCPNLRHVWKYNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SG  406
          |||| |: |  |||  |  | :|:||| |||||||  ||||| ::|:| |:       ||
Db   828  DL-EVIDIFKIEREGESQRYKPFRQLHNRLLWHGSRTNFAGILSGLRIAPPEAPVTG  886

Qy   407  GRVGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGF  466
          |  || |||:|:||| :|||:  ||| ::|||:|  |  |: | :    |||   |:| |
Db   887  YMFGKGIYFADMVSKSANYCHTSQ--GDPIGLIMLGEVALGNMYELK-HASHISKLPKGK  943

Qy   467  DSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRL  525
          |:|:  | | |:|  |:| :|     ::  ||:  ::  :| : ::|  |:: ||   :|
Db   944  HSVKGLGKTTPDPSASITLE---GVEVPLGTGI-PSGVNDTALLYNEYIVYDIAQVNL  997

Qy   526  RYLLEI  531
          ||:||:
Db   998  KYLLKL  1003
```

```
RESULT 12
PPOL_SARPE
ID   PPOL_SARPE    STANDARD;    PRT;    996 AA.
AC   Q11208;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
OS   Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7386;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=94170813; PubMed=8125121;
RA   Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,
RA   de Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;
RT   "Cloning and functional expression of poly(ADP-ribose) polymerase
RT   cDNA from Sarcophaga peregrina.";
RL   Eur. J. Biochem. 220:607-614(1994).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D16482; BAA03943.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_2N_FINGER_1; FALSE_NEG.
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   DNA_BIND      1    369       BY SIMILARITY.
FT   DOMAIN      370    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      382    456       BRCT.
FT   DOMAIN      508    996       NAD-BINDING.
FT   ZN_FING      19     54       BY SIMILARITY.
FT   ZN_FING     126    164       BY SIMILARITY.
FT   DOMAIN      211    214       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      232    235       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
SQ   SEQUENCE   996 AA;  113018 MW;  690DDD36E7487298 CRC64;

Query Match             22.9%;  Score 645.5;  DB 1;  Length 996;
Best Local Similarity   32.8%;  Pred. No. 4e-39;
Matches  175;  Conservative 101;  Mismatches 205;  Indels 53;  Gaps

Qy      22 EEEDSFRSTAEALRAAPADNRV-----IRVDPSCPF-------SRNPGIQVHEDYDCTL 68
             || || ||| :  ||  :       :|  ||  |||               :|| | :||:
Db     489 EESKSKSKSIYTKSVPKSMTLKIKDGLAVDPDSGLEDVAHYVSRN----KEKYNVVL 543

Qy      69 NQTNIGNNNKFYIIQLLEEG--SRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKKK 125
             |:|  |:| ||  ::|||| : :|| ::: :| |  ||::: :|  || |
Db     544 GITDIQKNKNSFYKLQLLESDMKNRFWVFRSWGRIGTTIGGNKLDNFSNLVDAIVQFKEL 603

Qy     126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLD 185
             |   |||:||   :   :|| ||  ::   :          :| || | |  : :|::|
Db     604 YLEKSGNHFENRENFVKVAGRMYPIDIDYAEDS--------KIDLSAEHDIKSKLPL- 652

Qy     186 PATQNLTNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 245
             : :|:| :|||||     :| | |||  ||||||::|:| |       || | |   ::|||
Db     653 -SVQDIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKIQSAYKVLTEIYELIQG-GGT 710

Qy     246 GQSLEELSSCFYTVIPIHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE 305
             : |: |||||||  |        ||      : | ||  || ||  |||  :|::|: ||
Db     711 NAKFIDATNRFYTLIPIHNFGTQSPPLLDTTEQVEQLRQMLDSLIEIECAYSLLQT--ED 767

Qy     306 EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKVN 362
             | : ||||  | :| :| : :  :  ::  || |||| |  ||  |  | |:  |||||
Db     768 SKAD--INPIDKHYEQLKTKLEPLDKNSEEYILLQKYVKNTHAETHKLYDLEVVDIFKVA 825

Qy     363 REGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFASE 418
             ||||| |  |::|:|  ||||||| | ||||||||||||||:||:   ||| |: ||| |||
Db     826 RQGEARRYKPFKLHNRRLLWHGSRLTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADM 885

Qy     419 NSKSAGYVTTMHCGHQVGYMFGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPD 478
             ||:|| :  :  | ||| | :| ||||  |::|  | | : |||
Db     886 VSKSANYCCTSH--HNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFGRGRTMPN 942

Qy     479 PAQDIELELDGQPVVVPQGPPVQCPSFKSSSSFSQSEYLLYKESQCRLRYLLEIH 532
             |    |  |   | ||   :  | :|      ||:::|:|||  | :| :  :||
Db     943 PSESIIRE-DG--VEIPLGKPITNDSLK-SSLLYNEFIIYDIAQVNIQYMLRMN 992
```

```
RESULT 13
PPOL_DROME
ID   PPOL_DROME    STANDARD;    PRT;    994 AA.
AC   P35875; Q9W5Q5; Q9W5S1;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   PARP OR CG17696/CG17718.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93234521; PubMed=8475096;
RA   Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,
RA   Sugimura T., Miwa M.;
RT   "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:
RT   leucine zipper in the auto-modification domain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND
RP   TISSUE SPECIFICITY.
RC   STRAIN=CANTON-S;
RX   MEDLINE=98234380; PubMed=9555614;
```

```
RA   Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;
RT   "Genomic organization of Drosophila poly(ADP-ribose) polymerase and
RT   distribution of its mRNA during development.";
RL   J. Biol. Chem. 273:11881-11886(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES
CC       OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM
CC       IN LATER EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN
CC       EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```

```
DR   EMBL; D13806; BAA02964.1; -.
DR   EMBL; AF051548; AAC24518.1; -.
DR   EMBL; AF051544; AAC24518.1; JOINED.
DR   EMBL; AF051545; AAC24518.1; JOINED.
DR   EMBL; AF051546; AAC24518.1; JOINED.
DR   EMBL; AF051547; AAC24518.1; JOINED.
DR   EMBL; AE002935; AAF45400.1; -.
DR   EMBL; AE002666; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE002892; AAF45445.2; ALT_SEQ.
DR   PIR; A47474; A47474.
DR   HSSP; P26446; 1A26.
DR   FlyBase; FBgn0010247; Parp.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP.reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.
FT   DNA_BIND      1    367
FT   DOMAIN      368    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      380    454       BRCT.
FT   DOMAIN      508    994       NAD-BINDING.
FT   ZN_FING      19     54       PARP-TYPE.
FT   ZN_FING     123    161       PARP-TYPE.
FT   DOMAIN      208    210       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   VARSPLIC    376    564       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   994 AA;  113791 MW;  ACA85A270DD29E08 CRC64;
```

```
Query Match             22.0%;  Score 619;  DB 1;  Length 994;
Best Local Similarity   31.58%;  Pred. No. 3.4e-37;
Matches  168;  Conservative 100;  Mismatches 222;  Indels 44;  Gaps 17;

Qy   16 KQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED----YDCT 67
           :  ::    ::           ::  :        :  |   |:|   :|
Db  481 KSRIPKETTKSLNSNSIYTKSMPV-SRTFKVKDGLAVDPDSGLEDIAHVYVDSNNKYSVV 539

Qy   68 LNQTIGNNNKFYIIQLL--EEGSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKK 124
          |:| |  ::|||   |::|   |||:|    |:|:|   ||:|:  ::|||
Db  540 LGLTDIQRNKNSYYKVQLLKADKKEKYWIFRSWGRIGTNIGNSKLEEFDYSESAKRNFKE 599

Qy  125 KFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSL 184
          : :| ||:::||:| || :   |||||:  :| :|||     :|  :|   |
Db  600 IYADKTGNEYEQRDNFVKRTGRMYPIEIQ--YDDDQKLVKHESHFFTS--------KL 647

Qy  185 DPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEEAMKNPTG 244
          ::  ||||:: |   || || ||||:::  ::| |::|:   |:| ::
Db  648 EISVQNLIKLIFDIDSMNKTLMEFHIDMDKMPLGKLSAHQIQSAYRVKEIYNVLECGSN 707

Qy  245 DGQSLEELSSCFYIVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEE 304
          | | ||:|::|  :: |||:|||   |   |:|||||| |:|||:|:|| ::|
Db  708 TAK-LIDATNRFYTLIPHNFGVQLPTLIETHQQIEDLRQMLDSLAEIEVAYSI----IK 761

Qy  305 EEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKV 361
          :|:   : ||:  || || :   ||  :|::|  :|::| :|  :|  ::||:|::|| :|:|
Db  762 SEDVSDACNPLDNHYAQIKTQLVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKV 821
```

```
Qy  362  NREGEGDRFQAHSKIGNRRLLWHGTNVAVVAAILTSGLRI----MPHSGGRVGKGIYFAS  417
         :::|| ||| :| | |  || ||||::||:|||  || ||||      :| |||||||||
Db  822  SRQGEARRFPFKKLHNRKLLWHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFAD  881

Qy  418  ENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHITIDDPSLKSPPPGFDSVIARGQTEP  477
         :|| ||::|||||| :|| ||::|| ||| :  :||:|||||||| ||  : :| : |
Db  882  MVSKSANYCCTSQ--QNSTGLMLLSEVALGDMMECT-SAKYINKLSNNKHSCFGRGRTMP  938

Qy  478  DPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEI  531
         || : | ||  | :  | |  |  || :  :| :|| :|| ||   :|:| |
Db  939  DPTKSY-IRSDG--VEIPYGETITDEHLK-SSLLYNEYIVYDVAQVNIQYLFRM  988

RESULT 14
YQN4_CAEEL
ID   YQN4_CAEEL     STANDARD;      PRT;    538 AA.
AC   Q09525;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.
GN   E02H1.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Smith A.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-
CC       RIBOSYLTRANSFERASE (EC 2.4.2.30).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z47075; CAA87379.1; -.
DR   HSSP; P26446; 1A26.
DR   WormPep; E02H1.4; CE01539.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   538 AA;  61268 MW;  3144E2S465FC7341 CRC64;

Query Match             14.1%;  Score 396.5;  DB 1;  Length 538;
Best Local Similarity   26.1%;  Pred. No. 2.1e-21;
Matches  149;  Conservative  87;  Mismatches 184;  Indels 151;  Gaps 24;

Qy   57  GIQVHEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFFCWNRWGRVGEVGQSKMNHFTCLE  116
         | ||      |  |:|| |||||||  :|| ||| :|  : || |  |  ::  :: |
Db   11  GYKVH------LCKTNIAQNNNKFYDMELLDEGGDFIVKLINGRIGYRGVTQLKFDDLD  64

Qy  117  DAKKDFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQRAVVKALSPQ---VDS  172
         :|||:| | |:|| :|:|||| :  ||:|  ::|      ::  ||   |      ||
Db   65  RAKKFFESKFYEKTHLHWEERDD-EPVPNKYAVVELATNARQTEEVKKEEPEPEKVDE  123

Qy  173  GPV----RTVVKPCSLDPATQNLITNIFSKEMFKNAMTL-------MNLDVKKM------  215
         ||:    |:  |       |: |  |:   |:  | ::|         | ::::
Db  124  KNTRGRKKRGIVKEKKEIKKEEPVEEV-NEKLKELMKCICDEDVHLGLLKQLKFNEAF  181

Qy  216  --PLGKLTKQQIARGFEALEALEEAMKNPT------------GDGQSLEELSS  254
             | |: :|  ||   |  | |             | |:::::|
Db  182  GRPIDCLSLAQLTTGYEILSKIEESIGGKSARRSTRGRPRVADRVLAVKSDGPSLHDINK  241
```

```
Qy  255  CFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVL----ADIELAQTLQAAPGEEEKVE  309
         :::|:|| |||: |||| :||  | ||:|:||:|       :| | ||   ||:::|
Db  242  -YYSLIPHSFGCVPPKIDSHAKIQAERELLDALKGSIEASLEKDLKKTASSKD----  295

Qy  310  EVPHPLDRDYQOLLRCColQLLDSGESEYKAIQTYLKQTGNSYRCPNLR---HVWKVN---  362
         | |   | : ::  | :| ::||| :||| | | ||| |||| |||:|   | |:|
Db  296  ----IYQRLYERLPCHLEPVSE---------EIAGKIGDCLAMRGPTHCYKLSLIDA  339

Qy  363  -------------------REGEGDRFQAHSKLG-------NRRLLWHGTNVAVVAALLTS  397
                            ::::| ||:|| ||:||        ||:|||||:|:|||:||:
Db  340  FELKDPNEIPTEAPVEVQEVPKKRGRKSTKTAAPTVPPPTKRLLWHGTRVTNVFSILMN  399

Qy  398  GLR--IMPHSGGRVGKGIYFASENSKSAGYVTTMHC--GGHQVGYMFLGEVALGK----  448
         ||:    | ||||||:|||||  |:|||:|| |:|    ||:|| ||:|||||| |
Db  400  GLQFPVGDRCGLMFGNGVYFANVPTKSANYC---CPEASKRV-FMLLCEVETANPLVLY  454

Qy  449  EHHITIDDPSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKS-  507
         | :| |||:|| |||  |||: :|  |:  :    : |  |   ::           :
Db  455  ESEIDADEKMEKAKK---TSVYAAGKHTPRDT---VEING---------IPAFKSN  495

Qy  508  -------SSFSQSEYLIYKESQCRLRYLLEI  531
                :::  |||::|:| |:|  :| ::|
Db  496  LETIEEETRLLYDEYVMFNKEHFKIKYVVEV  526

RESULT 15
PPOV_HUMAN
ID   PPOV_HUMAN     STANDARD;      PRT;   1724 AA.
AC   Q9UKK3; O75903; Q9H1M6; Q14682;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Vault poly(ADP-ribose) polymerase (EC 2.4.2.30) (VPARP) (193-kDa vault
DE   protein) (PARP-related/alphal-related H5/proline-rich) (PH5P).
GN   ADPRTL1 OR PARPL OR KIAA0177.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 306-319.
RX   MEDLINE=99408776; PubMed=1047748;
RA   Kickhoefer V.A., Siva A.C., Kedersha N.L., Inman E.M., Ruland C.,
RA   Streuli M., Rome L.H.;
RT   "The 193 kDa vault protein, VPARP, is a novel poly(ADP-ribose)
RT   polymerase.";
RL   J. Cell Biol. 146:917-928(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=20112770; PubMed=10644454;
RA   Still I.H., Vince P., Cowell J.K.;
RT   "Identification of a novel gene (ADPRTL1) encoding a potential
RT   poly(ADP-ribosyl)transferase protein.";
RL   Genomics 62:533-536(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Tromans A.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 94-1724 FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=96281124; PubMed=8724849;
RA   Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. v.
RT   The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 3:17-24(1996).
RN   [5]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=99198702; PubMed=10100603;
RA   Jean L., Risler J.-L., Nagase T., Couloarn C., Nomura N.,
RA   Salier J.-P.;
```

RT   "The nuclear protein PH5P of the inter-alpha-inhibitor superfamily: a
RT   missing link between poly(ADP-ribose)polymerase and the
RT   inter-alpha-inhibitor family and a novel actor of DNA repair?";
RL   FEBS Lett. 446:6-8(1999).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBUNIT: BINDS TO THE MAJOR VAULT PROTEIN (MPV). THE VAULT, A
CC   LARGE RIBONUCLEOPROTEIN COMPLEX, CONTAINS THE 100-kDA MPV AND 2
CC   PROTEINS OF 193 AND 240-kDA.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, BUT IS ALSO FOUND IN THE
CC   NUCLEUS, ASSOCIATED WITH MITOTIC SPINDLES.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; THE HIGHEST LEVELS ARE IN
CC   THE KIDNEY; ALSO DETECTED IN HEART, PLACENTA, LUNG, LIVER,
CC   SKELETAL MUSCLE, SPLEEN, LEUKOCYTES AND PANCREAS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF158255; AAD47250.1; -.
DR   EMBL; AF057160; AAC62491.1; -.
DR   EMBL; AL359763; CAC21562.1; -.
DR   EMBL; D79999; BAA11494.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR002035; vWFA.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00292; BRCT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation; Ribonucleoprotein.
FT   DOMAIN   1   94   BRCT.
FT   DOMAIN   256   569   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   876   1046   VWFA.
FT   DOMAIN   1562   1724   INTERACTION WITH THE MAJOR VAULT PROTEIN.
FT   DOMAIN   19   25   NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT   (POTENTIAL).
FT   DOMAIN   1237   1249   NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT   (POTENTIAL).
FT   CONFLICT   519   519   S -> P (IN REF. 1).
FT   CONFLICT   873   873   S -> N (IN REF. 2 AND 4).
FT   CONFLICT   897   897   Q -> E (IN REF. 2 AND 4).
FT   CONFLICT   899   899   A -> T (IN REF. 1).
FT   CONFLICT   936   936   A -> T (IN REF. 1).
FT   CONFLICT   936   936   A -> M (IN REF. 4).
FT   CONFLICT   1012   1012   V -> I (IN REF. 4).
FT   CONFLICT   1065   1065   A -> V (IN REF. 3 AND 4).
FT   CONFLICT   1080   1080   R -> L (IN REF. 3).
FT   CONFLICT   1108   1108   C -> R (IN REF. 3).
FT   CONFLICT   1265   1265   G -> A (IN REF. 2 AND 4).
FT   CONFLICT   1280   1280   G -> R (IN REF. 2 AND 4).
FT   CONFLICT   1328   1328   T -> P (IN REF. 3).
FT   CONFLICT   1331   1331   T -> A (IN REF. 3).
FT   CONFLICT   1394   1394   A -> S (IN REF. 3).
FT   CONFLICT   1459   1459   Y -> S (IN REF. 3).
FT   CONFLICT   1550   1550   P -> L (IN REF. 3).
FT   CONFLICT   1555   1555   V -> L (IN REF. 3).
FT   CONFLICT   1564   1564   T -> I (IN REF. 3).
FT   CONFLICT   1656   1656   P -> A (IN REF. 3).
SQ   SEQUENCE   1724 AA; 192587 MW; A009F3493460EDC CRC64;

Query Match   9.8%; Score 275; DB 1; Length 1724;
Best Local Similarity   21.7%; Pred. No. 7.4e-12;
Matches 121; Conservative 103; Mismatches 213; Indels 120; Gaps 20;

Qy   3 PKRKAS---VQTEGSKKQRQGTEEEDSFRSTAEALRAA-------PADNRVIRVDPSCPFSR 54
           | :||| |:||| | : : | ||:|:
Db   104 PDQKASSSEVKTEGLCPD-SATEEEDTVELTEFGMQNVEIPHLPQDFEVAKYNT----LE 158

Qy   55 NPGIQVHEDYDCTLNQTNIGNNNNKFYIIQ--LLEEG---SRFFCWNRWGRVGEVGQSKM 109
           |: :: |:| | ::.   : |::|| ||:|| |:|||
Db   159 KVGMEGGQEAVVVELQCSRDSRDCPFLISSHFLLDDGMETRRQFAIKKTS---------- 208

Qy   110 NHFTCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQ 169
           ||| |:|: : |: :| |:
Db   209 ------EDASEYFENYIEELKKQGFLLREHFTPE-------------------ATQLASEQ 244

Qy   170 VDSGPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGF 229
           :: :|: ::| :.   :|:   :::|:   :
Db   245 LQALLEEVMNSSTLSQEVSDLVEMIWAEALGHLEHMLLK-----PVNRISLNDVSKAE 298

Qy   230 EALEALEEAMKNPTGDGQSLEELSSC---FYTVIPHNFGKSRPPPINSPDVLQAKKDMLL 286
           : : ||:||| ||| :|| : : : | |!|
Db   299 GILLLVKAALKN----GETAEQLQKMMTEFYRLIPHK--GTMPKEVNL-GLLAKKADLCQ 351

Qy   287 VLAD-IELAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ 345
           :: :: : :| |||| |:::::
Db   352 LIRDMVNVCETNLSKPNP--------PSLAKYRALRCKIEHVEQNTEEFLRVRKEVLQ 401

Qy   346 TGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGL---RIM 402
           :| ||||| | |: |: : : ::| | || ||:| || : :::
Db   402 NHHSKSPVDVLQIFRVGRVNETTEFL--SKLGNVRPLLHGSPVQNIVGILCRGLLLLPKVV 459

Qy   403 PHSG------GRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 456
           | || | : ! |||| :::   : |:
Db   460 EDRGVQRTDVGNLGSGIYFSDSLSTSIKYSHPGETDGTRL--LLICDVALGKCMDLHEKD 517

Qy   457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYL 516
           || ||:||||:| : |: :|| |:
Db   518 FSLTEAPPGYDSVHGVSQT---------------------------ASVTTDFEDDEFV 549

Qy   517 IYKESQCRLRYLLEIHL 533
           :||: :|:||:::::::
Db   550 VYKTNQVKMKYIIKFSM 566

Search completed: August 29, 2002, 08:02:24
Job time: 301 sec

THIS PAGE BLANK (USPTO)

.Mon.Sep 9 08:38:42 2002

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: August 29, 2002, 08:01:34 ; Search time 117.38 Seconds
(without alignments)
785.537 Million cell updates/sec

Title: US-09-701-586B-8
Perfect score: 2813
Sequence: 1 MAPKRKASVQTEGSKKQRQG........EYLIYKESQCRLRYLLEIHL 533

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2772.5	98.6	528	11	Q91YR6	Q91yr6 mus musculu
2	2241	79.7	533	4	Q96CG2	Q96cg2 homo sapien
3	822.5	29.2	612	5	Q9TX06	Q9tx06 dictyosteli
4	717	25.5	653	10	O50017	O50017 zea mays (m
5	700.5	24.9	1014	11	Q921K2	Q921k2 mus musculu
6	686.5	24.4	607	13	Q9PS82	Q9ps82 gallus gall
7	686.5	24.4	607	13	Q9PS81	Q9ps81 xenopus. na
8	683.5	24.3	635	10	O81294	O81294 arabidopsis
9	621.5	22.1	983	10	Q9ZP54	Q9zp54 arabidopsis
10	621.5	22.1	1009	10	Q9SJW4	Q9sjw4 arabidopsis
11	610	21.7	593	5	Q9TX05	Q9tx05 drosophila
12	584.5	20.8	969	10	O24570	O24570 zea mays (m
13	584.5	20.8	980	10	Q92SV1	Q92sv1 zea mays (m
14	569.5	20.2	945	5	Q9N4H4	Q9n4h4 caenorhabdi
15	475.5	16.9	727	5	Q9XUA5	Q9xua5 caenorhabdi
16	310.5	11.0	2276	5	Q9TXQ1	Q9txq1 caenorhabdi

17	275.5	9.8	815	10	Q9SWB4	Q9swb4 glycine max
18	263.5	9.4	815	10	Q9FK91	Q9fk91 arabidopsis
19	156.5	5.6	1181	5	Q9X237	Q9x237 drosophila
20	156.5	5.6	1181	5	Q9VBP3	Q9vbp3 drosophila
21	146	5.2	1327	4	O95271	O95271 homo sapien
22	127.5	4.5	363	4	Q9H8R9	Q9h8r9 homo sapien
23	125.5	4.5	5198	5	O76518	O76518 caenorhabdi
24	124.5	4.4	1166	4	Q9H2K2	Q9h2k2 homo sapien
25	124.5	4.4	1265	4	Q9HAS4	Q9has4 homo sapien
26	120.5	4.3	181	12	O55721	O55721 chilo iride
27	119.5	4.2	954	10	P93826	P93826 arabidopsis
28	118.5	4.2	1092	2	O50236	O50236 zymomonas m
29	115	4.1	451	5	Q9V635	Q9v635 drosophila
30	114.5	4.1	1203	16	Q9CJI9	Q9cji9 lactococcus
31	112.5	4.0	935	10	Q9CA26	Q9ca26 arabidopsis
32	111.5	4.0	261	4	Q9H8F2	Q9h8f2 homo sapien
33	111.5	4.0	359	4	Q9Y4P7	Q9y4p7 homo sapien
34	111.5	4.0	1763	11	Q9JKX5	Q9jkx5 mus musculu
35	110	3.9	924	4	Q969W4	Q969w4 homo sapien
36	110	3.9	1223	4	Q9UFI5	Q9ufi5 homo sapien
37	108.5	3.9	1065	2	Q9AHK8	Q9ahk8 borrelia bu
38	108	3.8	949	16	Q92HW8	Q92hw8 rickettsia
39	108	3.8	1342	10	Q9FKN5	Q9fkn5 arabidopsis
40	107.5	3.8	920	5	Q9UIM9	Q9uim9 dictyosteli
41	107	3.8	1327	11	Q61595	Q61595 mus musculu
42	106.5	3.8	1412	4	Q96RT1	Q96rt1 homo sapien
43	106.5	3.8	1435	3	Q03291	Q03291 saccharomyc
44	106	3.8	757	16	Q9A0Q1	Q9a0q1 streptococc
45	106	3.8	769	10	Q9CG1	Q9cg1 arabidopsis

ALIGNMENTS

RESULT 1
Q91YR6
ID   Q91YR6   PRELIMINARY;   PRT;   528 AA.
AC   Q91YR6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 59.4 KDA PROTEIN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014870; AAH14870.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 528 AA; 59413 MW; 82EFB0C498EB5F74 CRC64;

Query Match   98.6%; Score 2772.5; DB 11; Length 528;
Best Local Similarity 98.9%; Pred. No. 1.3e-216;
Matches 527; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy   1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAPKRKASVQTEGSKKQRQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFFCWNRWCRVGEVGQSKMNHFTCLEDAKK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFFCWNRWCRVGEVGQSKMNHFTCLEDAKK 120

Qy 121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVK 180
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db 121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVK----VDSGPVRTVVK 175

Qy 181 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALELEEAMK 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   176  PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK  235
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Qy   241  NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAA   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   236  NPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELVQTLQAA   295

Qy   301  PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWK   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   296  PGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVWK   355

Qy   361  VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENS  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356  VNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASENS  415

Qy   421  KSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   416  KSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDPA  475

Qy   481  QDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL  533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   476  QDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL  528
```

RESULT   2
Q96CG2
ID   Q96CG2
AC   Q96CG2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 60.1 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014260; AAH14260.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   533 AA;   60089 MW;   6296A0E439CC7767 CRC64;

```
Query Match             79.7%;  Score 2241;  DB 4;  Length 533;
Best Local Similarity   80.1%;  Pred. No. 1.9e-173;
Matches   432;  Conservative   37;  Mismatches   58;  Indels   12;  Gaps   5;

Qy     1  MAPKRKASVQTEG--SKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI   58
          |||| ||  ||||  :|||| ||::|||||:|||||||:|::|||:||||||:|::|||
Db     1  MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT   60

Qy    59  QVHEDYDCTLNQTNIGNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED  117
          || |||:|||||||| ||||||||||| :: ||| |||||||||||||||||:||| |||
Db    61  QVYEDYNCTLNQTNIENNNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED  120

Qy   118  AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT  177
          |||||:||| :|||| | ||| ||::||||||||||||:|: :|:|| |||| | |||||
Db   121  AKKDFEKKFPREKTKNNWAERDHFVSHPGKYTLIEVQGAEDEAQEAVVK-----VDRGPVRT  175

Qy   178  V---VKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA  234
          |   |:|||||||||:|||||||||||||| |||:||||||||||||||||||||||||||
Db   176  VTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA  235

Qy   235  LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELA   294
          ||||:| ||:||||||||||:||||||||||:|:|||||:|:||||||||||||||||||
Db   236  LEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELA   295

Qy   295  QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPN  354
          |:|||  |::|:|:||||||||||||| :|||||||||::||||:||||::||:|:|||:
Db   296  QALQAV-SEQEKTVEEVPHPLDRDYQLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT   354
```

```
Db   355  LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIY  414
```
```
Qy   355  LRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIY  414
          |:|:|||||:||::||||||||||||:|||||||:|||||||||||||||||||||||||
Db   355  LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIY  414

Qy   415  FASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQ  474
          ||||||||||||  | |||:||||||||||||| |||||||:|:|||||||||||||||||
Db   415  FASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH  474

Qy   475  TEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL  533
          |||||:||:|||||||:||||||:||:||:|  |:|:|:|||||:|||||||||||||:||
Db   475  TEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL  533
```

RESULT   3
Q9TX06
ID   Q9TX06;
AC   Q9TX06;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26646; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
SQ   SEQUENCE   612 AA;   69241 MW;   CB340F7A88FF2364 CRC64;

```
Query Match             29.2%;  Score 822.5;  DB 5;  Length 612;
Best Local Similarity   38.3%;  Pred. No. 3.6e-58;
Matches   209;  Conservative   79;  Mismatches   207;  Indels   51;  Gaps   17;

Qy     6  KASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED--   63
          :| |  ||| :  : ||::|:::  :|| ||:::  |||||:      :|    :|:
Db    96  EAETKTAASDLDDSSSESEDEKNQISVKIKGRAAN-----DPHFPDSR--XHVYENGK   146

Qy    64  --YDCTLNQTNIGNNNKFYIIQLLE--EGSRFFCWNRWGRVGEVGQSKMNHF--TCLED  117
          | ||:||||| ||||:|:||||||  :|| ||||||||| | || |::     ||::
Db   147  DVYDATLNQTEISQNNNKXYIIQLLEADDGSSYSVWNRWGREGLKGQSSRKDFGKGGLNQ  206

Qy   118  AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT  177
          |  || |:| |:|||:  |  ::| || |||::|:  : ::|  ::       |  | ::
Db   207  AISLFCSKFYEKTKNFTTDRANFKKVAGKYDMIELDYSTDSKPK-NGASTATTTTKK   264

Qy   178  VV---KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEA  234
          ||   | ||||:|| |:::|:| ::   |:|:|:  |||||||||| | || :||:  :
Db   265  VVEHHKECSLDERVQELVKLIFDVKMMERTMTEAKYDLKKMPLGKLSKNQITKGYLVLKQ  324

Qy   235  LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELA   294
          :|:    : |      |:| ||||:|||:||| :.|||::  :| : :||: :|||| :|
Db   325  IEDVWGGKS--GESLSTLSSRFYTIIPHAFGMSVPPVINTNQMLIEKMNMLQNLADIEIA  382

Qy   295  QTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPN  354
          :|::           :|||||:: :|   :  : ::::: ||:  | ::  ||:  : |
Db   383  TNIIKDSESDESNI------LELHYAKLKTDIQPLDENSCEYKNILLYVKNTYQGGKKPT  436

Qy   355  LRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVG  410
          : ::|:::|:|::| ::   :||||:||||| : | | :|: ||||||||    ||  :|
Db   437  IVNIFKIDRDGEADRYKTKKHLGNRKLLWHGSRLTNYASIISQGLRIAPPEAPVSGYRFG  496
```

```
Qy  411 KGIYFASENSKSAGYVTTMHCGGHQVGY-MFLGEVALGKEHHITID---DPSLKSPPPGF 466
Db  497 KGVYFADCMSLSANYCRTV---GFLIDFCMLLGDVALGK----TADLARDTYMEKPQPNS 549
Qy  467 DSVIARGQTEPDPAQDIELELDGQPVVPQGPPVVQCPS-FKSSSFSQSEYLIYKESQCRL 525
Db  550 HSTWALGTVEPDPKVFFQ---DTEGCITPIYGQMI-PSQHKGVSCYEHQYVVYDVAQVHL 604
Qy  526 RYLLEI 531
Db  605 KYLLQL 610

RESULT  4
050017
ID   050017;          PRELIMINARY;       PRT;   653 AA.
AC   050017;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babylchuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants possess two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ222588; CAA10888.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
SQ   SEQUENCE  653 AA;  72995 MW;  5FD01923C4ABCD1D CRC64;

Query Match              25.5%;  Score 717;  DB 10;  Length 653;
Best Local Similarity    34.9%;  Pred. No. 1.5e-49;
Matches 190;  Conservative 91;  Mismatches 205;  Indels 58;  Gaps 17;

Qy   13 GSKKRQGTEEEDSFRS----TAEALRAAPADNRV---IRYDPSCPFSRNPGIQVHED-Y 64
Db  138 GAKEVIKGGDEVEVKKEKMVTATKKGAAVLDQHIPDHIKVNYHV------LQVGDEIY 190
Qy   65 DCTLNQTNIGNNNNKFYIIQLLEE--GSRFFCWNRWGRVGEVGQSKMNHFTCLED-AKKD 121
Db  191 DATLNQTNVGDNNNKFYIIQVLESDAGGSFMVYNRWGRVGRGQDKLHGPSTRDQAIYE 250
Qy  122 FKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQ-GEAESQEAVVKALSPQVDSGPVRTVVK 180
Db  251 FEGKFHNKTNNHWSDRKNFKCYAKKYTWLEMDYGETEK---------EIEKGSITDQIK 300
Qy  181 PCSLDPATQNLITNIFSKEMFKNAWTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 240
Db  301 ETKLETFRIAQFISLICNISMMKQRMVEIGYNAEKLPFGKLRKATILKGYHVLKRISDVIS 360
Qy  241 NPTGDGQSLEELSSCFYTVIPHNFG-RSRPPPINSPDVLQAKKDMLLVLADIELAQTLQ 298
Db  361 K--ADRRHLEQLTGEFYTVIPHDFGFRKMREFIIDFPQKLKAKLEMVEALGEIEIATKLL 418
Qy  299 AAPGEEEEKVEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRCPNL 355
Db  419 ------EDDSSDQDDPLYARYKQLHCDFTPLEADSDEYSMIKSYLRNTHGKTHSGYTVDI 472
```

```
Qy  356 RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK 411
Db  473 VQIFKVSRHGETERFQKFASTRNRMLLWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGK 532
Qy  412 GIYFASENSKSAGYVTTMHCGGHQGV--GYMFLGEVALGKEHHITIDDPSLKSPPPGFDS 468
Db  533 GVFADMFSKSANY-----CYASEACRSGVLLLCEVALGDMNELLNADYDANNLPKGKLR 587
Qy  469 VIARGQTEPDPAQDIELELDGQPVVVPQGPPVVQCPSFKSSSFSQSEYLIYKESQCRLRYL 528
Db  588 SKGVGQTAPNM---VESKVADDGVVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMRYV 643
Qy  529 LEIH 532
Db  644 LHVN 647

RESULT  5
Q921K2
ID   Q921K2;          PRELIMINARY;       PRT;   1014 AA.
AC   Q921K2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE)
DE   POLYMERASE).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC012041; AAH12041.1; -.
KW   Transferase.
SQ   SEQUENCE  1014 AA;  112721 MW;  5DCE68E4CB3F46EB CRC64;

Query Match              24.9%;  Score 700.5;  DB 11;  Length 1014;
Best Local Similarity    33.3%;  Pred. No. 6.2e-48;
Matches 182;  Conservative 100;  Mismatches 215;  Indels 49;  Gaps 17;

Qy    1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
Db  494 VAPGKSAAPSKKSK---GAVKEEGVNKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 549
Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
Db  550 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVIGSNKLEQMPSKED 609
Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
Db  610 AVEHFMKLLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEERAVKK------L 655
Qy  178 VVKP---CSLDPATQNLITNIFSKEMFKNAWTLMNLDVKKMPLGKLTKQQIARGFEALEA 234
Db  656 TVKPGTKSKLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQAAYSILSE 715
Qy  235 LEEAMKNPTGDGQSLEELSSCFYTVIPHNFGSRPPPINSPDVLQAKKDMLLVLADIELA 294
Db  716 VQQAVSQGSSESQIL-DLSNRFYTLIPHDFGFGMKKPPLLNNADSVQAKVEMLDNLLDIEVA 774
Qy  295 QTLQAAPGEEEEKVEVPHPLDRDYQLLRCQLQLLDSGSEYKAIQTYLKQ----TGNSY 350
Db  775 YSILRGGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAY 829
Qy  351 RCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SG 406
Db  830 DL-EVIDIFKIEREGESQRYKPFQLHNRRLLWHGSRTNFAGILSQGLRIAPPEAPVTG 888
Qy  407 GRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDDPSLKSPPPGF 466
```

```
Db  889 YMFGKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGK 945
        ::::  ||||        ||  :||||||:
Qy  467 DSVIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPS-FKSSSFSQSEYLIYKESQCRL 525
        :|| :  |||        |  ||| :  ||||| |:  ||||:|
Db  946 HSVKGLGKTTPDPSASITLE----GVEVPLGTGI--PSGVNDTCLLYNEYIYDIAQVNL 999

Qy  526 RYLLEI 531
        :|||::
Db 1000 KYLLKL 1005


RESULT  6
Q9PS82
ID  Q9PS82       PRELIMINARY;     PRT;    607 AA.
AC  Q9PS82;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96007847; PubMed=7578427;
RA  Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA  Kofler B., Schweiger M., Wagner E.F.;
RT  "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT  ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT  inactivation of the ADPRT gene in the mouse.";
RL  Biochimie 77:444-449(1995).
DR  HSSP; P26446; 1A26.
DR  InterPro; IPR001357; BRCT.
DR  InterPro; IPR001290; PARP.
DR  InterPro; IPR004102; PARP_reg.
DR  Pfam; PF00533; BRCT; 1.
DR  Pfam; PF00644; PARP; 1.
DR  Pfam; PF02877; PARP_reg; 1.
DR  PROSITE; PS50172; BRCT; 1.
SQ  SEQUENCE 607 AA; 68033 MW; 75F6EEID30D8F402 CRC64;


Query Match            24.4%;   Score 686.5;   DB 13;   Length 607;
Best Local Similarity  33.8%;   Pred. No. 3.9e-47;
Matches   184;  Conservative   97;  Mismatches 207;  Indels  57;  Gaps  20;

Qy    5 RKASVQTEGSKKQRQG-TEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHED 63
        |  ::::| |:| || :| || |: || : || :| | |:
Db   93 KPANMKSAGVKEEQGPSKSEKKMKLTVKGGAAVDPDSGL--EDSAHVFEKGGKI----- 145
Qy   64 YDCTLNQTNIGNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKK 120
          : | ||| :|| : ||||| |:   | ||| :  ||||| :  || |   | ||
Db  146 FSATLGLVDIVKGTNSYYKLQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAVE 205
Qy  121 DFKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVK 180
        :||  :| | :: |   :: | ||| |: :   |
Db  206 HFLNLYEEKTGNSWHSKN-FTKYPKKFYPLEID-YGQDEEAVRKL---TVSAGTKSKLAK 260
Qy  181 PCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMK 240
        |       :|  ::     |         |||:|  :||| :||: ||:
Db  261 P------IQDLLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILNEVQQAVS 314
Qy  241 NPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAA 300
            :| |||:    |  :  |    :      :|  : :|  :
Db  315 DGGSESQIL-DLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRG 373
Qy  301 PGEEEEKVEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ---TGNSYRCPNLR 356
        ||: :|   : ::| |: :: : || |||: :|:|  :|  :
Db  374 GNEDGDK-----DPIDINYEKLRTDIKVDKDSEEAKIIKQYVKNTHAATHNAYDL-KVV 427
Qy  357 HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKG 412
```

```
Db  428 EIFRIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKG 487
        ::::  ||||        ||  :||||||:
Qy  413 IYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEH------HITIDDPSLKSPPPGF 466
        |||  ||||| |        : : |||||||
Db  488 IYFRDMVSKSANYCHTSQ-ADPIGLILLGEVALGNMYELKNASHIT------KLPKGK 538
Qy  467 DSVIARGQTEPDPAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLR 526
        :|| :  |||        |  ||| :  ||||| |:  ||||:|
Db  539 HSVKGLGKTAPDPT--ATTTLDG--VEVPLGNGIS-TGINDTCLLYNEYIYDVAQVNLK 593

Qy  527 YLLEI 531
        :||::
Db  594 YLLKL 598


RESULT  7
Q9PS81
ID  Q9PS81       PRELIMINARY;     PRT;    607 AA.
AC  Q9PS81;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS  Xenopus.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC  Xenopodinae;
OX  NCBI_TaxID=8353;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96007847; PubMed=7578427;
RA  Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA  Kofler B., Schweiger M., Wagner E.F.;
RT  "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT  ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT  inactivation of the ADPRT gene in the mouse.";
RL  Biochimie 77:444-449(1995).
DR  HSSP; P26446; 1A26.
DR  InterPro; IPR001357; BRCT.
DR  InterPro; IPR001290; PARP.
DR  InterPro; IPR004102; PARP_reg.
DR  Pfam; PF00533; BRCT; 1.
DR  Pfam; PF00644; PARP; 1.
DR  Pfam; PF02877; PARP_reg; 1.
DR  PROSITE; PS50172; BRCT; 1.
SQ  SEQUENCE 607 AA; 67496 MW; 54CDEBBE22079886 CRC64;


Query Match            24.4%;   Score 686.5;   DB 13;   Length 607;
Best Local Similarity  34.1%;   Pred. No. 3.9e-47;
Matches   182;  Conservative   95;  Mismatches 212;  Indels  45;  Gaps  19;

Qy   10 QTEGSKKQRQGT-EEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHEDYDCTL 68
           :::||:| ||:: || :||:: |   :       |        :|
Db   97 KSSGVKEEKGSNKSEKKMKLTVKGGAAIDPDS---ELEDSCHVLETGG----KIFSATL 149
Qy   69 NQTNIGNNNKFYIIQLLE--EGSRFFCWNRWGRVGEV-GQSKMNHFTCLEDAKKDFKKK 125
           :|| :|| : ||||| :   | ||| :  ||||| :  || |   | |||
Db  150 GLVDITRGTNSYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKEDAIEHFLNL 209
Qy  126 FWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLD 185
           |   |      :|| |   :: | ||| |:  :   |
Db  210 YQDKTGNAWHS-PNFTKYPKKFYPLEI--DYGQEEDVVKKLS--VGAGTKSKLAKP---- 260
Qy  186 PATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD 245
             |:|| |     ::     |         |||:|  :||| :||:
Db  261 --VQELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAVSESLSE 318
Qy  246 GQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEE 305
           :| |||:    |  :  |    :      :|  :
Db  319 ARLL-DLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDG 377
```

```
Matches  190;  Conservative  90;  Mismatches  220;  Indels  51;  Gaps  19;

Qy    2  APKRKASVQTEGSKKQRQGTEEEDSFRS--TAEALRAAPADNRVIRVDPSCPFSRNPGIQ   59
         ||..| :: : ||   |  |||||||            |  ||
Db  109  APVKSSNDEAE---DDNNGFEEEKKEEKIVTATVKGAAVLDQWI----PDEIKSQYHVLQ  161

Qy   60  VHED-YDCTLNQTNIGNNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMN-HFTCL  115
          :|| ||||||||.||||:|||||:|:|||||    |||||||:|||.||||| |||:
Db  162  RGDDVYDAILNQTNVRDNNNKFVLQVLESDSKKTYMVYTRWGRVGVKGQSKLDGPYDSW  221

Qy  116  EDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPV  175
          |: :|| :| : | : | : | |
Db  222  DRAIEIFTNKFNDKTKNYWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNNDIPSSSS---  278

Qy  176  RTVVKP--CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALE  233
          .| :     ||  |: .:::| :: |: .    |:|||| ||||| |.|: .|.|.|||
Db  279  --EVKPEQSKLDTRVAKFISLICNVSMMAQHMEIGYNANKLPLGKISKSTISKGYEVLK   336

Qy  234  ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPP--INSPDVLQAKKDMLLVLADI  291
          .: : :  . .  |||||| :.||||||| || :  :|  || | :|.|: :::.| .|
Db  337  RISEVIDR--YDRTRLEELSGEEYTVIPHDFGFKKMSQFVIDFPQKLKQKIEMVEALGEI  394

Qy  292  ELA-QTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---G  347
          |||  .| :  | ::|       ||| |||.|||||||||::||| ||||::||.    |
Db  395  ELATKLLSVDPGLQDD-------PLYYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKT  447

Qy  348  NSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH--  404
          :| |: .|  . ..| | :|||||.|. ||.::|||||:.:.| |.|.|||| ||:
Db  448  HSGYTVEIAQLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAP  507

Qy  405  -SGGRVGKGIYFASENSKSAGYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPSLK  460
           .| |.|||:||| .. |.: |   || |.   | ||||| |||||| :.::  ||  |
Db  508  VTGYMFGKGVYFADMFSKSANY-----CYANTGANDGVLLCEVALGDMNELLYSDYNAD  562

Qy  461  SPPPGFDSVIARGQTEPDAQDIELELDGQPVVPQGPPVQCPSFKSSSSFSQSEYLIYKE  520
          . || | |. . | .  : :  || ||  ||| |.|| . : :|  :..| :||:||:
Db  563  NLPPGKLSTKGVGKTAPNPSEAQTLE-DG--VVVPLGKPVE-RSCSKGMLLYNEYIVYNV  618

Qy  521  SQCRLRYLLEI  531
          : :||:.:|:
Db  619  EQIKMRYVIQV  629


RESULT  9
Q92P54

ID   Q92P54        PRELIMINARY;        PRT;      983 AA.
AC   Q92P54;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).
GN   PARP-1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LANDSBERG ERECTA;
RA   Doucet-chabeaud G., Kazmaier M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ131705; CAA10482.1; -.
DR   HSSP; P26446; 1A26.
KW   Transferase; Glycosyltransferase; NAD.
FT   CHAIN         2    983       POLY(ADP-RIBOSE) POLYMERASE.
SQ   SEQUENCE   983 AA;  111232 MW;  468E12A8EF1B6F4F CRC64;

Query Match             22.1%;  Score  621.5;  DB  10;  Length  983;
Best Local Similarity   32.7%;  Pred. No. 1.5e-41;
  Matches  179;  Conservative  87;  Mismatches  209;  Indels  73;  Gaps  21;
```

```
Qy  306  EKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLK---QTGNSYRCPNLRHVWKV  361
         ||     |||| |   : :  |  ::  :||  ||.:::    :| || .|||||..|:
Db  378  EK-----DPIDVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNAYDLEVL-EIFKI  431

Qy  362  NREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFAS  417
          ||||  || .  :| |::|||||:. .| |.|.|:|||| ||      | ||||||||
Db  432  DREGEYQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFAD  491

Qy  418  ENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEP  477
          . :|||.|   |   |.  ||||||||||:: | :  |  :|.       | :
Db  492  MVSKSANYCHAM--PGSPIGLILLGEVALGNHELKAASQITKL-PKGKHSVKGLGRTAP  548

Qy  478  DPAQDIELELDGQPVVPQGPPVQCPSFKSSSSFSQSEYLIYKESQCRLRYLLEI  531
          || :   ||.|   | | |.   .|  ||  |.| :|:||.: :..:|:||:|
Db  549  DPSATV--QLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLKYLLKL  597


RESULT  8
O81294

ID   O81294        PRELIMINARY;        PRT;      635 AA.
AC   O81294;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE).
GN   T14P8.19 OR AT4G02390.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Waterston R.;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RA   Kalicki J., Elliott G., Cloud J.;
RT   "The sequence of A. thaliana T14P8.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RP   EU Arabidopsis sequencing project;
RA   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF069298; AAC19283.1; -.
DR   EMBL; AL161494; CAB80732.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
KW   Transferase.
SQ   SEQUENCE   635 AA;   72017 MW;  E3F1CBE4D367A377 CRC64;

Query Match             24.3%;  Score  683.5;  DB  10;  Length  635;
Best Local Similarity   34.5%;  Pred. No. 7.3e-47;
```

```
Qy    15 KKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH----ED----YDC 66
         ||||  ::     :: :  : |::|| :  :  ::| |::     | ||  :
Db   472 KKQRKLPFDKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT 524

Qy    67 TLNQTNIGNNNNKFYIIQLLEE--GSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDF 122
         || :   :  :::  |:: :::::  : :| | |||||| : :::::: :   ::|
Db   525 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGGNKVEEMS-KSDAVHEF 583

Qy   123 KKKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVV 179
         | : ||||:    | ::   :|  :::| |  :: :  : :    ::
Db   584 KRLFLEKTGNTWESWEQKTNFQKQPGKFLPLDIDYGVNKQVA--------KKEPFQT- 632

Qy   180 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL----EAL 235
          :   || :  |  ::| | :  ::  : |  ::|::::|||||| :: |:|||    |
Db   633 -SSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQRLL 691

Qy   236 EEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQ 295
          :: : | :  ::|: :  ::|:||   |  |  | :   :::: :: ::||  ||:: :
Db   692 TESDPQPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIEIAS 747

Qy   296 TLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP-- 353
          : :   |  :    :  | :  ::: |  :: |  ::: |::::  |   |: : : |
Db   748 RI---VGFDVDSTES---LDDKYKKLLHCDISPLPHDSEDYRLIEKYL----NTTHAPTH 796

Qy   354 -----NLRHVWKVNREGEGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRI----MP 403
              : |  : | ||||| ||: || ||||:::||||| :::  : ||::||||    |
Db   797 TEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAP 856

Qy   404 HSGGRVGKGIYFASENSKSAGYVTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPP 463
          :|  :|||||||:  :||||:| ::  | ::|: ||| :|||| :|| | ||  |: ||
Db   857 ATGYMFGKGIYFADLVSKSAQYCYT--CKKNPVGLMLLSEVALGEIHELT-KAKYMDKPP 913

Qy   464 PGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQC 523
          | :|  | | :   :  ||: |  |||:|| |  : | :  | | :  : || ||:::
Db   914 RGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPVS-SKVKASELMYNEYIVYDTAQV 969

Qy   524 RLRYLLEI 531
          :|::|| :
Db   970 KLQFLLKV 977
```

RESULT 10
Q9SJW4
ID   Q9SJW4   PRELIMINARY;   PRT;   1009 AA.
AC   Q9SJW4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE POLY (ADP-RIBOSE) POLYMERASE.
GN   At2g31320.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X.;

RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006593; AAD20677.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; Znf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE   1009 AA; 114133 MW; CDE6E41CC2A3A2DB CRC64;

Query Match   22.1%; Score 621.5; DB 10; Length 1009;
Best Local Similarity   32.7%; Pred. No. 1.6e-41;
Matches 179; Conservative   87; Mismatches 209; Indels   73; Gaps   21;

```
Qy    15 KKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVH----ED----YDC 66
         ||||  ::     :: :  : |::|| :  :  ::| |::     | ||  :
Db   498 KKQRKLPFDKYKIEDTSESLVTVKVKGRSAVHEAS------GLQEHCHILEDGNSIYNT 550

Qy    67 TLNQTNIGNNNNKFYIIQLLEE--GSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDF 122
         || :   :  :::  |:: :::::  : :| | |||||| : :::::: :   ::|
Db   551 TLSMSDLSTGINSYYILQIIQEDKGSDCYVFRKWGRVGNEKIGGNKVEEMS-KSDAVHEF 609

Qy   123 KKKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVV 179
         | : ||||:    | ::   :|  :::| |  :: :  : :    ::
Db   610 KRLFLEKTGNTWESWEQKTNFQKQPGKFLPLDIDYGVNKQVA--------KKEPFQT- 658

Qy   180 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEAL----EAL 235
          :   || :  |  ::| | :  ::  : |  ::|::::|||||| :: |:|||    |
Db   659 -SSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQRLL 717

Qy   236 EEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQ 295
          :: : | :  ::|: :  ::|:||   |  |  | :   :::: :: ::||  ||:: :
Db   718 TESDPQPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIEIAS 773

Qy   296 TLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP-- 353
          : :   |  :    :  | :  ::: |  :: |  ::: |::::  |   |: : : |
Db   774 RI---VGFDVDSTES---LDDKYKKLLHCDISPLPHDSEDYRLIEKYL----NTTHAPTH 822

Qy   354 -----NLRHVWKVNREGEGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRI----MP 403
              : |  : | ||||| ||: || ||||:::||||| :::  : ||::||||    |
Db   823 TEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAP 882

Qy   404 HSGGRVGKGIYFASENSKSAGYVTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPP 463
          :|  :|||||||:  :||||:| ::  | ::|: ||| :|||| :|| | ||  |: ||
Db   883 ATGYMFGKGIYFADLVSKSAQYCYT--CKKNPVGLMLLSEVALGEIHELT-KAKYMDKPP 939

Qy   464 PGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQC 523
          | :|  | | :   :  ||: |  |||:|| |  : | :  | | :  : || ||:::
Db   940 RGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPVS-SKVKASELMYNEYIVYDTAQV 995

Qy   524 RLRYLLEI 531
          :|::|| :
Db   996 KLQFLLKV 1003
```

RESULT 11
Q9TX05
ID   Q9TX05   PRELIMINARY;   PRT;   593 AA.
AC   Q9TX05;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Drosophila sp. (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7242;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   593 AA;   68018 MW;   D9BA37E38B8E7CCD CRC64;

Query Match   21.7%;
Best Local Similarity   31.3%;   Score 610; DB 5; Length 593;
Matches 167; Conservative 100; Mismatches 223; Indels 44; Gaps 17;

```
Qy    16  KQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQ------VHED----YDCT   67
              ::  :| |                 |:|  |:|          |:|     |:|
Db    80  KSRIPKETTKSLNSNSIYTKSMPV-SRFFKVKDGLAVDPDSGLEDIAHVYVDSNNKYSVV  138

Qy    68  LNQTNIGNNNKFYIIQLL--EEGSRFFCWNRWGRVG-EVGQSKMNHFTCLEDAKKDFKK   124
              | ::| ||:|:|:|   :| |::     |:  |  |  |:|:||||||:|::|:||
Db   139  LGLTDIQRNKNSYYKVQLLKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKE  198

Qy   125  KFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSL  184
              :|:|||: |    |||||||||||||||||:||::: :||| |:::
Db   199  IYADKTGNEYEQRDNFVKRTGRMYPIEIQ--YDDDQKLVKHESHFFTS--------KL  246

Qy   185  DPATQNLITNIFSKEMFKNAMTLMNLDVKKMPGKLTKQQIARGFEALEALEEAMKNPTG  244
              |:::|:| |     ||: : :| ||| :|||||| :   :: ::
Db   247  EISVQNLIKLIFDIDSMNKTLMEFHIDMDKMPGKLSAHQIQSAYRVRVKEIYNVLECGSN  306

Qy   245  DGQSLEEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEE  304
              :    ||| : :  |:|||| |: |   ||||| |: :|||| ||||||| |  |||
Db   307  TAK-LIDATNRFYTLIPHNFGVQLPTLIETHQQIEDLRQMLDSLAEIEVAYSI-----IK  360

Qy   305  EEKVEEVPHPLDRDYQLLRCQLQLDSGESEYKAIQTYLKQT-GNSYRCPNLR--HVWKV  361
              |: ||:  |: :||:||:  |::: ||  |: :| |  |||:  |:|:
Db   361  SEDVSDACNPLDNHYAQIKTQLVALDKNSEEFSILSQVKNTHASTHKSYDLKIVDVFKV  420

Qy   362  NREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRI---MPHSGGRVGKGIYFAS  417
              :|:|    :| |  | ||||: |:|: :: |  ||||| ::    |:| |||||||||
Db   421  SRQGEARREKPFKLHNRKLLWHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFAD  480

Qy   418  ENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEP  477
              |::|||||||| ||||:||:|||||||||||:| :|:||||||::|||| | :|||:|:|
Db   481  MVSKSANYCCTSQ--QNSTGLMLLSEVALGDMMECT-SAKYINKLSNNKHSCFGRGRTMP  537

Qy   478  DPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFQSEYLIYKESQCRLRYLLEI    531
              |:::::::::  :| ::   || :::: :|::|||: ::|:||:  || :::
Db   538  DPTKSY-IRSDG--VEIPYGETITDEHLK-SSLLYNEYIVYDVAQVNIQYLFRM  587
```

RESULT 12
O24570
ID   O24570   PRELIMINARY;   PRT;   969 AA.
AC   O24570;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.

OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babiychuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants possess two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ222589; CAA10889.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; 2nf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 2.
SQ   SEQUENCE 969 AA; 109128 MW; EB23AC62EEC14009 CRC64;

Query Match   20.8%;
Best Local Similarity   30.7%;   Score 584.5; DB 10; Length 969;
Matches 164; Conservative 98; Mismatches 199; Indels 73; Gaps 20;

```
Qy    33  ALRAAPADNRVIRVDPSCPFSRNPGIQ-----VHED----YDCTLNQTNIGNNNKFYIIQ   84
              ||  : :|      |                 :|||           :::: ||::|
Db   468  ALESSKGSTVTVKVKGRSAVHESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQ  527

Qy    85  LLE--EGSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKKDFKKFWEKTKNKWEE---R  137
              ||:   | ::   :|: |     |:|||:: |  ||:||::|::|:|:|:|:|:
Db   528  IIEQDDGSECYVFRKWGRVGSEKIGGQKLEMSKTE-AIKEFKRLFLEKTGNSWEAWECK  586

Qy   138  DRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLDPATQNLITNIFS  197
              :: |   |  | |:  :  : ::     |           :: |  :
Db   587  TNFRKQPGRFYPLDV------DYGVKKAPKRKDISEMKS-----SLAPQLLELMKMLFN  634

Qy   198  KEMFKNAMTLMNLDVKKMPGKLTKQQIARGFEALEALEEAMKNPTGDGQSLEE-----L  252
              | |:::| |: | :|::  |:||: |::|:::| |:|      ||::
Db   635  VETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAA  694

Qy   253  SSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEKVEEVP  312
              :||:|::|||| |:|||  ||||| |  :::|   :|: : || :   : |::|::
Db   695  SNRFFTLIP------SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDES----  745

Qy   313  HPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------NLRHVWKVNREG  365
              | ||    :  |  :| :::  |:|:::| :|: |:|||          |:|::|:|:|
Db   746  --LDDKYMKLHCDITPLAHDSEDYKLIEQYLLNT----HAPTHKDWSLEEEVFSLDRDG  799

Qy   366  EGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFASENS  420
              |||| |  |:|||||||||||:|: :: |  ||||:|||    |||||:|||||||::
Db   800  ELNKYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVS  859

Qy   421  KSAGYVTTMHC--GGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEP  477
              |||||||| ||   | ::::::::||||||::: :|:||||| ||| :||::|||:|||
Db   860  KSAQY-----CYVDRNNPVGLMLLSEVALGDMYELK-KATSMDKPPRGKHSTKGLGKTVP  913

Qy   478  DPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFQSEYLIYKESQCRLRYLLEI    531
              ||:::::::::|:  :   || ::| :|::|:|| |||:::: :::|:
Db   914  LESEFVKWRDD---VVPCCGKPVP-SSIRSSELMYNEYIVYNTSQVKMQFLLKV    963
```

RESULT 13
Q9ZSV1
ID   Q9ZSV1   PRELIMINARY;   PRT;   980 AA.

```
AC   Q9ZSV1;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP)-RIBOSE POLYMERASE (EC 2.4.2.30).
GN   PARP1.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99026291; PubMed=9808734;
RA   Mahajan P.B., Zuo Z.;
RT   "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL   Plant Physiol. 118:895-905(1998).
DR   EMBL; AF093627; AAC79704.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR001510; znf-PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase.
SQ   SEQUENCE  980 AA;  110475 MW;  9D8AED26BC37E5C1 CRC64;
```

```
Query Match             20.8%;  Score 584.5;  DB 10;  Length 980;
Best Local Similarity   30.7%;  Pred. No. 1.6e-38;
Matches  164;  Conservative  98;  Mismatches 199;  Indels  73;  Gaps  20;

Qy    33 ALRAAPADNRVIRVDPSCPFSRNPGIQ----VHED----YDCTLNQTNIGNNNNKFYIIQ  84
         |  ||                  :|||||  :  ||::: ::   |:: ||||::|:
Db   479 ALESSKGSTVTVKVKGRSAVHESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQ 538

Qy    85 LLE--EGSRFFCWNRWGRVG--EVGQSKMNHFTCLEDAKDFKKKFWEKTKNKWEE---R 137
         ::| :: :|| |||||  :||||| ||| : | |  :  |||:|||||  ||
Db   539 IIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEKTGNSWEAWECK 597

Qy   138 DRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVKPCSLDPATQNLITNIFS 197
         :: |          |:| ||  ::|   ||   ::: ::   :  :::
Db   598 TNFRKQPGRFYPLDV------DYGVKKAPKRKDISEMKS-----SLAPQLLELMKMLFN 645

Qy   198 KEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSLEE-----L 252
         ||  ||||::|| :||:|||| |:| | :   :: ||  :: :|
Db   646 VETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAA 705

Qy   253 SSCFYTVIPHFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEEKVEEVP 312
         | :|:|::| |||||::|||| || :|| |: ||:||||||:|| |: ||  ::
Db   706 SNRFFTLIP---SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDES----- 756

Qy   313 HPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCP------NLRHVWKVNREG 365
         ||| | |||  :||  :|| ||| | | || | |: ||:|
Db   757 --LDDKYMKLHCDITPLAHDSEDYKLIEQYLLNT---HAPTHKDWSLEEEVFSLDRDG 810

Qy   366 EGDRFQAH-SKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFASENS 420
         || | | ::|||  ||||: | |  :::|  ||  ||:
Db   811 ELNKYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVS 870

Qy   421 KSAGYVTTMHC---GHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEP 477
         |||:            |||:| :|| || |  |:  |||  | || |   |:||||| |
Db   871 KSAQY----CYVDRNNPVGLMLLSEVALGDMYELK-KATSMDKPPRGKHSTKGLGKTVP 924

Qy   478 DPAQDIELELDGQPVVVPQGPVQCPSFKSSSFSQSEYLLYKESQCRLRYLLEI 531
```

```
Db   925 LESEFVKWRDD---VVVPCGKPVP-SSIRSSELMYNEYIVYNTSQVKMQFLLKV 974
             ::::        |||| ||||::|  :| ::||        :|| :::||:||||::::||::
```

```
RESULT 14
Q9N4H4
ID   Q9N4H4              PRELIMINARY;       PRT;    945 AA.
AC   Q9N4H4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 108.0 KDA PROTEIN.
GN   Y71F9AL.18.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Bradshaw-Cordum H., Scott K., Graves T.;
RT   "The sequence of C. elegans cosmid Y71F9AL.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024200; AAF36011.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; znf-PARP; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE  945 AA;  108006 MW;  1D0A62C954BC6AD9 CRC64;
```

```
Query Match             20.2%;  Score 569.5;  DB 5;  Length 945;
Best Local Similarity   31.5%;  Pred. No 2.4e-37;
Matches  156;  Conservative  80;  Mismatches 201;  Indels  59;  Gaps  14;

Qy    64 YDCTLNQTNIGNNNNKFYIIQLLEEGSR--FFCWNRWGRVG-EVGQSKMNHFTCLEDAKK 120
         |||:||||:| : ||||::|||||::|::   |||||||||| :||||||||||:| :||
Db   477 YQATLSFTDLTQNKNSYYKIQLLKDDQRENYVFRSWGRVGTEVGGNKHESYSNSNEAIL 536

Qy   121 DFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRTVVK 180
         || ||| ||||:|  |:  :|    :|| :  |  ::  :  |:
Db   537 KFQDVFHEKTKNDWIYRKHFRKMPGMFSYVETDYSEFAQ---------ITDTEIT 582

Qy   181 PCS---LDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
         ||  :   |  | |:|||::|||||||||:|||||||||||||||||  :  :
Db   583 PGSKTLLPKSVKEVVMSIFDVENMKSALKSFEMDVNKMPLGRLSHNQINLAFEVLNDISD 642

Qy   238 AMKNPTGDGQSLEELSSCFYTVIPHFGRSRPPINSPDVLQAKKDMLLVLADIELAQTL 297
         |:||| |||| ::| :|||| :|||||||||||::||||::||||||:|||:||||||
Db   643 LLVKLPIDASRILDFSNKFYTIIPHFNFGMRVPEPIDSFHKIKEKNMNLNALLDIKFAYD- 701

Qy   298 QAAPGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCP-NL 355
         |  |:        ::|  |:   :|  :   ||: :| |:|:  : ::| :||    ||
Db   702 QISGGDVPASTSLGIDPVDINYQKLKCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVEL 761
```

```
Qy   356  RHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGK  411
          |:|:| :|: | :||||||||::   |:|||||| ||  ||  ||
Db   762  IDILKLNRDNESSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPPEAPVSGYMFGK  819

Qy   412  GIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFDSVI  470
          |:|| |||  :   : |:|||  |||||||||||||||||||| :|:|:|||||||:|
Db   820  GVYFADMFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSVQ  876

Qy   471  ARGQ--------TEPD---PAQDIELELDGQPVVPQGPPVQCPSFKSSSFSQSEYLI  517
          |:       :| |:|   |
Db   877  GLGRQCPREIGSYNHPDGYTIPLGLTYMQLQGKQDV---------DYHLLYNEFIV  923

Qy   518  YKESQCRLRYLLEIHL  533
          | | |::||:|:|:
Db   924  YDVDQIQLKYLVRVKM  939
```

```
RESULT 15
Q9XUA5
ID   Q9XUA5      PRELIMINARY;      PRT;    727 AA.
AC   Q9XUA5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC8.1 PROTEIN.
GN   AC8.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McMurray A.A.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z83097; CAB05448.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; 2nf-PARP; 1.
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 1.
SQ   SEQUENCE 727 AA; 82884 MW; 530ABA8E991FFEFD CRC64;
```

```
Query Match             16.9%;   Score 475.5;  DB 5;  Length 727;
Best Local Similarity   30.2%;   Pred. No. 7e-30;
Matches 134;  Conservative  68;  Mismatches 170;  Indels  71;  Gaps  13;

Qy   126  FWEKTKNKWEERDRFVAQPNKYTLIEVQGE--AESQEAVVKALSPQVDSGPVRTVVKPCS  183
          |:|         ::                   : :|        |
Db   315  FHEKTKNDWIYRKHFRKMPGMFSVVETDYSEFVGNNGHKKKITPGSKTLLPKS  374

Qy   184  LDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQOIARGFEALEALEEAMKNPT  243
          :   ::: :|| |:||||||::|||||:||| |:::| : |::|  ::   |
Db   375  V----KEVVMSIFDVENMKSALKSFEIDVNKMPGRLSHNQINLAFEVLNDISDLLVKLP  430

Qy   244  GDGQSLEELSSCFVTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGE  303
          : ::::||:||||||||  | ::|: :|    :||| | ::|:||   | ||
Db   431  IDASKILDFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNMLNALLDIKFAYD-QISGGD  489

Qy   304  EEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQYLKQT-GNSYRCP-NLRHVWKV  361
          : :: ::  |  |:  :: : :|  :: ::  :   |  |::   :  :|  | |
Db   490  VPASTSLSIDPVDINYRKLKCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVELIDILKV  549
```

```
Qy   362  NREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGGRVGKGIYFAS  417
          ||: :|: :||||||||::   |:|||||| ||  ||  ||:
Db   550  NRDNESSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPPEAPVSGYMFGKGVYFAD  607

Qy   418  ENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFDSVIARGQTE  476
          |:|||  :   : |:|||  |||||||||||||||||| :|:|:|||||||:|  | ||
Db   608  MFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSV-------  657

Qy   477  PDPAQDIELELDGQPVVPQGPPVQCP----SFKSS---------SFSQ--  512
          |: :       :| |:|   |      ::          :|
Db   658  --------------QGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQGKQNVDYHL  698

Qy   513  --SEYLIYKESQCRLRYLLEIHL  533
          ::|:| | | |::||:|:|:
Db   699  LYNEFIVDVDQIQLKYLVRVKM  721
```

Search completed: August 29, 2002, 08:01:37
Job time: 369 sec

us-09-701-586b-8.rspt

Mon Sep 9 08:38:42 2002

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:       August 29, 2002, 07:59:24 ; Search time 69.02 Seconds
                                            (without alignments)
                                            742.040 Million cell updates/sec

Title:        US-09-701-586B-4
Perfect score: 2823
Sequence:     1 MAPKPKPWVQTEGPEKKKGR..........EYLIYQESQCRLRYLLEVHL 533

Scoring table: BLOSUM62
              Gapop 10.0 , Gapext 0.5

Searched:     283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters:     283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    PIR_71:*
              1:  pir1:*
              2:  pir2:*
              3:  pir3:*
              4:  pir4:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	2419	85.7	459	2	T08713	NAD+ ADP-ribosyltr
2	694	24.6	635	2	T01311	NAD+ ADP-ribosyltr
3	677	24.0	1016	1	JS0428	NAD+ ADP-ribosyltr
4	674.5	23.9	996	1	S42208	NAD+ ADP-ribosyltr
5	670.5	23.8	653	2	T03656	probable NAD+ ADP-
6	669.5	23.7	1011	1	JH0581	NAD+ ADP-ribosyltr
7	666.5	23.6	1014	1	A29725	NAD+ ADP-ribosyltr
8	657	23.3	500	2	S26057	NAD+ ADP-ribosyltr
9	653	23.1	998	2	S31735	NAD+ ADP-ribosyltr
10	649.5	23.0	1013	1	S04200	NAD+ ADP-ribosyltr
11	638	22.6	994	1	A47474	NAD+ ADP-ribosyltr
12	598.5	21.2	983	2	T51353	NAD+ ADP-ribosyltr
13	598.5	21.2	1009	2	C84719	probable poly (ADP
14	578	20.5	969	2	T03657	NAD+ ADP-ribosyltr
15	534	18.9	727	2	T18600	hypothetical prote
16	393.5	13.9	538	2	T20414	NAD+ ADP-ribosyltr
17	190.5	6.7	135	2	PN0494	NAD+ ADP-ribosyltr
18	164.5	5.8	2004	2	D88948	protein zK1005.1 [
19	134	4.7	1156	2	B70356	chromosome assembl
20	126	4.5	1222	2	C88504	protein B0361.3 [i
21	120.5	4.3	2044	2	AB1180	probable peptidogl
22	120	4.3	1938	1	A40997	myosin heavy chain
23	116.5	4.1	1435	2	S69632	regulatory protein
24	113.5	4.0	181	2	T03058	NAD+ ADP-ribosyltr
25	113.5	4.0	1134	2	A60234	IgA Fc receptor pr
26	113.5	4.0	1164	1	FCSOAG	IgA Fc receptor pr
27	113.5	4.0	4540	2	T30838	cytoplasmic dynein
28	110.5	3.9	2078	2	T25400	hypothetical prote
29	110	3.9	1534	2	A56734	ribosome receptor,
30	110	3.9	2288	2	T29999	hypothetical prote
31	109.5	3.9	446	1	A46335	gag polyprotein -
32	108.5	3.8	388	2	QOYV	transforming prote
33	108.5	3.8	1233	2	T30534	chromosome segrega
34	108.5	3.8	1871	2	S27938	hypothetical prote
35	108	3.8	359	2	T12540	hypothetical prote
36	108	3.8	1092	2	T33717	carbamoyl-phosphat
37	108	3.8	3225	2	I52300	giantin - human
38	108	3.8	3259	2	A56539	giantin - human
39	107.5	3.8	716	1	T26998	hypothetical prote
40	107	3.8	1186	2	G69708	chromosome segrega
41	106.5	3.8	568	1	T34522	hypothetical prote
42	106	3.8	725	1	JC5016	hyaluronan recepto
43	106	3.8	932	2	S62555	protoplast regener
44	106	3.8	1199	2	T29145	hypothetical prote
45	106	3.8	2176	2	T13806	toucan gene protei

ALIGNMENTS

```
RESULT       1
T08713
NAD+ ADP-ribosyltransferase homolog DKFzp566G0224.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Nov-2000
C;Accession: T08713
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16472
A;Accession: T08713
A;Molecule type: mRNA
A;Residues: 1-459 <ANS>
A;Experimental source: fetal kidney; clone DKFzp566G0224
C;Genetics:
A;Note: DKFzp566G0224.1
C;Cross-references: EMBL:AL050034
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match         85.7%;  Score 2419; DB 2;  Length 459;
Best Local Similarity 99.8%; Pred. No. 6.3e-162;
Matches  458; Conservative   1; Mismatches   1; Indels   0;  Gaps   0;

Qy    75  IENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREKTK 134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  IENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREKTK  60

Qy   135  NNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLIT 194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  NNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLIT 120

Qy   195  NIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELS 254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELS 180

Qy   255  SHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHP 314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  SHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHP 240

Qy   315  LDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVNQEGEEDRFQAHS 374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  LDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVNQEGEEDRFQAHS 300

Qy   375  KLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENSKSAGYVIGMKCGAH 434
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   301  KLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGEGIYFASENSKSAGYVIGMKCGAH 360

Qy   435  HVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELEDGQQVVV 494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  HVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELEDGQQVVV 420

Qy   495  PQGQVPCPEFSSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
```

```
Db   421  ||||||||||||||||||||||||||||||||||
          PQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 459

RESULT   2
T01311
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C;Accession: T01311; S65662
R;Kalicki, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of A. thaliana T14P8.
A;Experimental source: cultivar Columbia
R;Lepiniec, L.; Babiychuk, E.; Kushnir, S.; Van Montagu, M.; Inze, D.
FEBS Lett. 364, 103-108, 1995
A;Title: Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(ADP-r
A;Reference number: S65662; MUID:95269779
A;Accession: S65662
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-115,'Gt',116-635 <LEP>
A;Cross-references: EMBL:Z48243; NID:g853721; PIDN:CAA88288.1; PID:g853722
C;Genetics:
A;Gene: PARP
A;Map position: 4
A;Introns: 28/3; 116/1; 162/3; 190/1; 268/3; 312/1; 330/1; 355/2; 371/1; 387/3; 408/3; 4
A;Note: T14P8.19
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger

  Query Match      24.6%;  Score 694;  DB 2;  Length 635;
  Best Local Similarity  34.2%;  Pred. No. 9.3e-41;
  Matches  185;  Conservative  97;  Mismatches  221;  Indels  38;  Gaps  18;

Qy    11  TEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRV-DPTCPLSSNPGTQVYED--- 65
          |:||:| ||| :| |   | ||| ||| :|| :|| ||  | |:|:||| ||| |
Db   107  SNAPVKSSNDEAEDDNNGFEEEKKEKIVTATKKGAAVLDQWIPDEIKSQYHVLQRGDDV 166

Qy    66  YNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK 123
          |  |||| | |||||||||| ||| | || |||||||||| || |::|  ||| ||||
Db   167  YDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVYTRWGRVGKGQSKLDGPYDSWDRAIE 226

Qy   124  DFEKKFREKIKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQP- 182
          ||:||| | |:||| |:||||||||||| |:||  ||:|| | |  |  | |:|| ||
Db   227  IFTNKFNDKTKNYWSDRKEFIPHPKSYTWLEMDYGKEENDSPVND---IPSSSSEVKPE 283

Qy   183  -CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 241
           |||| | ||| |: ||:|| ||| | ||||||||||| |:|| |:|:| |  |  ||
Db   284  QSKLDTRVAKFISLICNVSMMAQHMEIGYNANKLPLGKISKSTISKGYEVLKRISEVI- 342

Qy   242  GPTDGGQSLEELSSHFYTVIPHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ 299
          | ||| :| | || ||| |: ||||||||    |||| | || | ||| |||| || ||
Db   343  -DRYDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLL 401

Qy   300  AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ- 356
          |:| || || |  | |: |  ||  ||:| ::|||| | |:|||| : | |  |  |:
Db   402  SVDPGLQ-----DDPLYYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGVTVEIA 456

Qy   357  HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKG 412
          ||| |::|  | :   |:|||| | :|| | |  |  ||| || | |    ||| ||||
Db   457  QLFRASRAVEADRFQQFSSSSKNRMLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKG 516

Qy   413  IYFASENSKSAGYVIGMKCCAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGDSVIAR 472
          ||| | |:| ||| | |   |||||| | || |||| || || |||:| |:| ||| |
Db   517  VYFADMFSKSANYCYA-NTGAND-GVLLLCEVALGDMNELLYSDYNADNLPPGKLSTKGV 574
```

```
Db   473  GHTEPDPTQDTELELDGQQVVVPQGQPV--PCPEFSSSTFSQSEYLIYQESQCRLRYLLE 530
          | | ||:|||||||::|||  ||||||::  ||||||||||||||||::||||||||||
Db   575  GKTAPNPSEAQTLE-DG--VVVPLGKPVERSC---SKGMLLYNEYIVNVEQIKMRYVIQ 628

Qy   531  V 531
          |
Db   629  V 629

RESULT   3
JS0428
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine
N;Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: JS0428; S00328; A30458
R;Saito, I.
submitted to JIPID, February 1990
A;Reference number: JS0428
A;Accession: JS0428
A;Molecule type: mRNA
A;Residues: 1-1016 <SAI>
A;Experimental source: thymus
R;Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; T
Eur. J. Biochem. 171, 571-575, 1988
A;Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int
A;Reference number: S00328; MUID:88151954
A;Accession: S00328
A;Molecule type: mRNA
A;Residues: 648-714; 838-904 <TAN>
A;Cross-references: EMBL:X06986
A;Accession: A30458
A;Molecule type: protein
A;Residues: 658-685;689-696;893-901 <TA2>
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-loo
F;21-51/Region: zinc finger
F;128-165/Region: zinc finger
F;200-220/Region: helix-turn-helix motif
F;224-231/Region: nuclear location signal
F;250-270/Region: helix-turn-helix motif
F;494-501/Region: nucleotide-binding motif A (P-loop)
F;890-903/Region: nucleotide binding #status predicted

  Query Match      24.0%;  Score 677;  DB 1;  Length 1016;
  Best Local Similarity  33.8%;  Pred. No. 2.8e-39;
  Matches  187;  Conservative  99;  Mismatches  205;  Indels  62;  Gaps  24;

Qy    1   MAPKPK----PWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSS 56
          || ||     |  :||||||||||||||||||||||||||||| | :||||:||||| ||
Db   495  VGPKGKSGAAPSKKSKGPVKEG--TNKSEKRMKLTLKGGAA-------VDPDSGLEH 543

Qy    57  NPGTQVYED---YNCTLNQTNIENNNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV-GQS 110
          ||| |      |||||| ||  |||||| || |:|| |:|| |||||||||||  | ||
Db   544  N--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLEDDKESRYWIFRSWGRVGTVIGSN 601

Qy   111  KINHFTRLEDAKKDFEKKFREKTKNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVD 169
          ||:||||||||  | |||||| ||||| |||||||||||| |:|  |  || |:||| |
Db   602  KLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKHPKKFYPLEIDYGQDE--EAVKKLT 658

Qy   170  RGPVRVTKRVQPCSLDPATQKLITNISKEMFKNTMALMDLDVKKMPLGKLSKQQIARG 229
          |||||| || ||: |  |:|| || |:|:|  || |:|| |  | |:|||| |:| |
Db   659  VNP-GTKSKLPKP------VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAA 711

Qy   230  FEALEALEEAL-KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLV 288
          | |||| | |  | || || :|:::   | ||| ||| ||| ||||:| | |||| |||
Db   712  YSILSEVQQALSQGSSD--SHILDLSNRFYTLIPHDFGMKPPLLNNANSVQAKVEMLDN 769

Qy   289  LADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLLDSGAPEYKVIQTYLEQT-G 347
          |||||| | || |:|  :    | | |     ||  ||:| ||||| ||:|| | : |
Db   770  LLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDKDSEEAEIIRKYVKNTHA 825
```

```
Qy  348  SNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH- 404
            :  |||||| |||| | | ||||    ||||||
Db  826  TTHNAYDLEVVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPE 885

Qy  405  ---SGGRVGKGIYFASENSKSAGYVIGMKCGAHH--VGYMFLGEVALGREHHINTDNPS 458
            ||||| | |  ||           || ||
Db  886  APVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEAALGNMYELKHAR-H 939

Qy  459  LKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIY 518
            :  ||:|    ||     :  |:||
Db  940  ISKLPKGKHSVKGLGKTPDPS--ASITVDG--VEVPLGTGI-SSGVNDTCLLYNEYIVY 994

Qy  519  QESQCRLRYLLEV 531
            :  :|:|:
Db  995  DIAQVHLKYLLKL 1007
```

RESULT 4
S42208
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - flesh fly (Sarcophaga peregrina)
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Sarcophaga peregrina
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S42208; S71496
R;Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurat
Eur. J. Biochem. 220, 607-614, 1994
A;Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from Sarc
A;Reference number: S42208; MUID:94170813
A;Accession: S42208
A;Molecule type: mRNA
A;Residues: 1-996 <MAS>
A;Cross-references: EMBL:D16482; NID:g473742; PIDN:BAA03943.1; PID:g538248
A;Accession: S71496
A;Molecule type: protein
A;Residues: 170-188;721-736;813-819;879-885 <MAX>
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger
F;1-369/Domain: DNA binding #status predicted <DNA>
F;370-507/Domain: auto-modification #status predicted <AMO>
F;508-996/Domain: NAD binding #status predicted <NAD>

```
Query Match             23.9%;  Score 674.5;  DB 1;  Length 996;
Best Local Similarity   33.6%;  Pred. No. 4.1e-39;
Matches  179;  Conservative  104;  Mismatches  195;  Indels  55;  Gaps  22;

Qy   25  EEDPFRSTAEALKAIPAEKRI----IRVDPTCPLSSNPGTQVY------EDYNCTLNQTN 74
            : :  : |  |                :        ||||||
Db  490  ESKSSKSKSIYTKSVPKSMTLKIKDGLAVDPDSGLED--VAHVVSRNKEKYNVVLGITD 547

Qy   75  IENNNKFYIIQLLQD--SNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFRE 131
            | :  || |||| :  :  || | |||||||  :   | ||   ||  |||||
Db  548  IQKNKNSFYKLQLLESDMKNRFWV-FRSWGRIGTTIGGNKLDNFSNLVDAIVQFKELYLE 606

Qy  132  KTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQ 190
            | :  :  |   :    |                  ||
Db  607  KSGNHFENRENFVKVAGRMYPIDIDYAEDS------KIDLSAEHDIKSKL-PLSV---Q 655

Qy  191  KLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEAEALKGPTDGGQSL 250
            : | :  ||  :  :  ||||| | |||||||||:  :       |   ||||||
Db  656  DIIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLTEIYELIQG--GGTNA 712

Qy  251  E--ELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQAQAVSEQEKTV 308
            |    : :   |||||||| | ||      :       : :  |
Db  713  KFIDATNRFYTLIPHNFGTQSPPLLDTTEQVEQLRQMLDSLIEIECAYSLLQTEDSKADI 772

Qy  309  EEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEG 365
            :  : :|     ||      :         :       ||:|:|:
Db  773  ----NPIDKHYEQLKTKLEPLDKNSEEYILLQKVVKNTHAETHKLYDLEVVDIFKVARQG 828

Qy  366  EEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSK 421
            | |||||||||||||||||||      |||||||||||
```

```
Db  829  EARRYKPFKKLHNRRLLWHGSRLTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADMVSK 888

Qy  422  SAGYVIGMKCGAHH--VGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP 479
            || |        |   |||||||||||
Db  889  SANYC---CTSHHNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFGRGRTMPNP 943

Qy  480  TQDTELEDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVH 532
            :  :  ||      :    |:  ::  ::        :   |    |  |
Db  944  SESIIRE-DG--VEIPLGKPITNDSLKSSLL-YNEFIIYDIAQVNIQYMLRMN 992
```

RESULT 5
T03656
probable NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Oct-1999
C;Accession: T03656
R;Babylchuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Mo
submitted to the EMBL Data Library, November 1997
A;Description: Higher plants possess two poly(ADP-ribose) polymerases.
A;Reference number: Z14991
A;Accession: T03656
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-653 <BAB>
A;Cross-references: EMBL:AJ222588; NID:e1264090; PIDN:CAA10888.1; PID:e1264091
C;Genetics:
A;Gene: PARP
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

```
Query Match             23.8%;  Score 670.5;  DB 2;  Length 653;
Best Local Similarity   35.4%;  Pred. No. 4.3e-39;
Matches  170;  Conservative  93;  Mismatches  186;  Indels  31;  Gaps  14;

Qy   64  EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKINHFTRLED-A 121
            |  ||||| | || || |||| |:|  | :  :  || || ||     |: |: ||  |
Db  188  EIYDATLNQTNVGDNNKFYIIQVLESDAGGSFMVYNRWGRVGRGQDKLHGPSPTRDQA 247

Qy  122  KKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ 181
            :  :  ||  || || | |  : :   | ::  : |            :    :   ::
Db  248  IYEFEGKFHNKTNNHWSDRKNFKCYAKKYTWLEMDYGETEKE----IEKG---SITDQIK 300

Qy  182  PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 241
              : |     :| | |  :  : |  : |     ||||||  :  | | :
Db  301  ETKLETRIAQFISLICNISMMKQRMVEIGYNAEKLPLGKLRKATILKGYHVLKRISDVIS 360

Qy  242  GPTDGGQSLEELSSHFYTVIPHNFGHSQPP--INSPELLQAKKDMLLVLADIELAQALQ 299
                  :  |  | :||||||:||    :   :: |  | |:   :   | |||| | |
Db  361  --KADRRHLEQLTGEFYTVIPHDFGFRKMREFIIDTPQKLKAKLEMVEALGEIEIATKL- 417

Qy  300  AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ-- 356
                 :      | :|     |  ||    :   |  : : :  |  : |::|:|
Db  418  ---LEDDSSDQDDPLYARYKQLHCDFTPLEADSDEYSMIKSYLRNTHGKTHGKTHSGYTYDIV 473

Qy  357  HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGKG 412
            :  |: :  || :||| |  ||||||||| :  |||||||| | ||
Db  474  QIFKVSRHGETERFQKFASTRNRMLLWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGKG 533

Qy  413  IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 472
            :||| :  ||| |    |      | |  |||||  :    ||   |  |   | | :
Db  534  VYFADMFSKSANYCYASE--ACRSGVLLCEVALGDMNELLNADYDANNLPKGKLRSKGV 591

Qy  473  GHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVH 532
            | | |  :  :       |||||| | |    :  :  |:  |:||  |    |  | :
Db  592  GQTAPNMV-ESKVADDG--VVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMRYVLHVN 647
```

RESULT 6
JH0581
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - chicken
N;Alternate names: poly(ADP-ribose) synthase
C;Species: Gallus gallus (chicken)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JH0581
R;Ittel, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.
Gene 102, 157-164, 1991
A;Title: Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence and c
A;Reference number: JH0581; MUID:91340148
A;Accession: JH0581
A;Molecule type: mRNA
A;Residues: 1-1011 <ITT>
A;Cross-references: EMBL:X52690; NID:g63742; PIDN:CAA36917.1; PID:g63743
C;Comment: This protein is a chromatin-bound enzyme.
C;Comment: This enzyme catalyzes DNA-dependent post-translational modifications of varic
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc fi

Query Match   23.7%;   Score 669.5; DB 1;   Length 1011;
Best_Local_Similarity 33.2%;   Pred. No. 9.5e-39;
Matches 186; Conservative 100; Mismatches 205; Indels 69; Gaps   24;

Qy   7 PW--VQTEGPE-------------KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPT 51
             : | :| . ||: : | | | |. ||:: :
Db   477 PWGAEVKTEHQEVAVDGKCSKPANMKSAGKVKEQGPSKSERKMKLTV---KGGAAVDPD 533

Qy   52 CPLSSNPGTQVYED----YNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWGRVG 105
             : | || .| : | || | : | ||| : : ||||| |
Db   534 SGLEDS--AHVFEKGGKIFSATIGLVDIVKGTNSYYKIQLLEDDRESRYWV-FRSWGRVG 590

Qy   106 EV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQE 163
             :| |:| | | | |: | : : | | || | || || | ||| : :|
Db   591 TVIGSNKLEQMPSKEDAVEHFLNLYEEKTGNSWHSK-NFTKYPKKFYPLEIDYGQDEAV 649

Qy   164 AVVKDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSK 223
             :   . : | : | | . | | || :| :|||| : | :|||||||:||||||
Db   650 RKLTVSAG---TKSKLAKP------IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSK 700

Qy   224 QQIARGFEALEALEEALKGPTDGGQSLE--ELSSHFYTVIPHNFGHSQPPINSPELLQA 281
             :| | | : | ||:   :|| ||: | | : | | || :| | :|
Db   701 RQIQSAYSILNEVQQAV---SDGGSESOILDLSNRFYTLIPHDFGMKKPPLLSNLEYIQA 757

Qy   282 KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYOLLKCQLQLLDSGAPEYKVIQT 341
             | | | | :| | | | |:| | | | | || | || | ||| | | .:
Db   758 KVQMLDNLLDIEVAYSLLRGGNEDGDKD----PIDINYEKLRTDIKVVDKDSEEAKIIKQ 813

Qy   342 YLEQT-GSNHRCPTLQ--HIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSG 398
             ||| : | | | : | | | | : ||||| | |||: | | ||
Db   814 YVKNTHAATHNAYDLKVVEIFRIEREGESQRYKPFKQLHRQLLWHGSRTTNFAGILSQG 873

Qy   399 LRIMPH-----SGGRVGKGIYFASENSKSAGYVIGMKC---GAHHVGYMFLGEVALGREHH 451
             |||| |||| | | | | | | || || ||| | || :|
Db   874 LRIAPPEAPVTGYMFGKGIYFRDMVSKSANY-----CHTSQADPIGLILLGEVALGNMYE 928

Qy   452 INTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFS 511
             | || | | ||| | || | | || | | | | | || | | ||||: :|
Db   929 LKNASHITKL-PKGKHSVKGLGKTAPDPTATT--TLDG--VEVPLGNGI-STGINDTCLL 982

Qy   512 QSEYLIYQESQCRLRYLLEV 531
             : |||| || |||:|||::
Db   983 YNEIVYDVAQVNLKYLLKL 1002

RESULT   7
A29725
NAD+ ADP-ribosyltransferase (EC 2.4.2.30), nuclear - human
N;Alternate names: poly (ADP-ribose) polymerase; poly (ADP-ribose) synthetase; poly(ADP)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
A;Accession: A29725; A28498; A39976; A26901; I38096; B33321; A33321; A35635; A61559; S14
R;Uchida, K.; Morita, T.; Sato, T.; Ogura, T.; Yamashita, R.; Noguchi, S.; Suzuki, H.; N
Biochem. Biophys. Res. Commun. 148, 617-622, 1987
A;Title: Nucleotide sequence of a full-length cDNA for human fibroblast poly(ADP-ribose)
A;Reference number: A29725; MUID:88076933
A;Accession: A29725

A;Molecule type: mRNA
A;Residues: 1-69,'Q',71-1014 <UCH>
A;Cross-references: GB:M18112; NID:g190166; PIDN:AAA60137.1; PID:g190167
R;Kurosaki, T.; Ushiro, H.; Mitsuuchi, Y.; Suzuki, S.; Matsuda, M.; Matsuda, Y.; Katu
J. Biol. Chem. 262, 15990-15997, 1987
A;Title: Primary structure of human poly (ADP-ribose) synthetase as deduced from cDNA
A;Reference number: A28498; MUID:8058958
A;Accession: A28498
A;Molecule type: mRNA
A;Residues: 1-16,'E',18-211,'K',213-236,'R',238-356,'H',369-1014 <KUR>
A;Cross-references: GB:J03473
R;Cherney, B.W.; McBride, O.W.; Chen, D.; Alkhatib, H.; Bhatia, K.; Hensley, P.; Smul
Proc. Natl. Acad. Sci. U.S.A. 84, 8370-8374, 1987
A;Title: cDNA sequence, protein structure, and chromosomal location of the human gene
A;Reference number: A39976; MUID:88068596
A;Accession: A39976
A;Molecule type: mRNA
A;Residues: 1-49,'D',51-612,'Q',614-907,'Y',909-939,'R',941-979,'I',981-1014 <CHE>
A;Cross-references: GB:J03030
A;Note: the authors translated the codon ATA for residue 980 as Asn
R;Suzuki, H.; Uchida, K.; Shina, H.; Sato, T.; Okamoto, T.; Kimura, T.; Miwa, M.
Biochem. Biophys. Res. Commun. 146, 403-409, 1987
A;Title: Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and expressi
A;Reference number: A26901; MUID:87298455
A;Accession: A26901
A;Molecule type: mRNA
A;Residues: 441-610,'N',612-880;921-1014 <SUZ>
A;Note: the sequence figure has an omission of forty residues
R;Ogura, T.; Nyunoya, H.; Takahashi-Masutani, M.; Miwa, M.; Sugimura, T.; Esumi, H.
Biochem. Biophys. Res. Commun. 167, 701-710, 1990
A;Title: Characterization of a putative promoter region of the human poly(ADP-ribose)
A;Reference number: I38096; MUID:90211250
A;Accession: I38096
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-40 <RES>
A;Cross-references: EMBL:X16674; NID:g510112; PIDN:CAA34663.1; PID:g1017423
R;Auer, B.; Nagl, U.; Herzog, H.; Schneider, R.; Schweiger, M.
DNA 8, 575-580, 1989
A;Title: Human nuclear NAD(+) ADP-ribosyltransferase(polymerizing): organization of t
A;Reference number: A33321; MUID:90091744
A;Accession: B33321
A;Molecule type: DNA
A;Residues: 38-43;93-98;132-137;204-209;237-242;276-281;335-340;384-389;431-436;512-5
A;Cross-references: GB:M29544; GB:M22953
A;Note: the authors translated the codon GTG for residue 54 as Glu
A;Note: these fragments represent intron-exon boundaries
A;Accession: A33321
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 16-66;96;121-159,'D',161-167 <AU2>
A;Note: these fragments represent a zinc finger-containing DNA-binding region
R;Gradwohl, G.; Menissier de Murcia, J.; Molinete, M.; Simonin, F.; Koken, M.; Hoeijm
Proc. Natl. Acad. Sci. U.S.A. 87, 2990-2994, 1990
A;Title: The second zinc-finger domain of poly(ADP-ribose) polymerase determines spec
A;Reference number: A35635; MUID:90222155
A;Accession: A35635
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 12-26,'T',28-66;116-166 <GRA>
R;Schneider, R.; Auer, B.; Kuehne, C.; Herzog, H.; Klocker, H.; Burtscher, H.J.; Hirs
Eur. J. Cell Biol. 44, 302-307, 1987
A;Title: Isolation of a cDNA clone for human NAD (+): protein ADP-ribosyltransferase.
A;Reference number: A61559; MUID:88082900
A;Accession: A61559
A;Molecule type: mRNA
A;Residues: 381-420;682-710 <SCH>
R;Yokoyama, Y.; Kawamoto, T.; Mitsuuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Tera
Eur. J. Biochem. 194, 521-526, 1990
A;Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.
A;Reference number: S14010; MUID:91099327
A;Accession: S14010
A;Status: preliminary

```
A;Molecule type: DNA
A;Residues: 1-95 <YOK>
A;Cross-references: EMBL:X56140; NID:g35286; PIDN:CAA39606.1; PID:g825702
C;Genetics:
A;Gene: GDB:ADPRT; PPOL
A;Cross-references: GDB:119508; OMIM:173870
A;Map position: 1q41-1q42
C;Comment: This protein can ADP-ribosylate itself as well as other proteins.
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransfer

Query Match    23.6%;  Score 666.5;  DB 1;  Length 1014;
Best Local Similarity    33.4%;  Pred. No. 1.5e-38;
Matches 186;  Conservative 101;  Mismatches 207;  Indels 63;  Gaps 24;

Qy     7 PW-------VQTEGPEKKKGRQAGR------EEDPFRSTAEALKAIPAEKRIIRVDPTC 52
Db   480 PWGAEVKAEPVEVVAPRGKSGAALSKKSKGQVKEEGINKSEKRMKL--TLKGGAAVDPDS 537

Qy    53 PLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQD-SNRFFTCWNRWGRVGE 106
Db   538 GLEHS-AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYW-IFRSWGRVGT 594

Qy   107 V-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEA 164
Db   595 VIGSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EA 651

Qy   165 VVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQ 224
Db   652 VKKLTVNP-GTKSKLPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKR 704

Qy   225 QIARGFEALEALEAALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKD 284
Db   705 QIQAAYSILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVE 763

Qy   285 MLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLE 344
Db   764 MLDNLLDIEVAYSLLRGGSDDSSKD---PIDVNYEKLKTDIKVVDRDSEEAEIIRKYVK 819

Qy   345 QT-GSNHRCPTLQ-HIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRI 401
Db   820 NTHATTHNAYDLEVIDFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRI 879

Qy   402 MPH----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINT 454
Db   880 APPEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEVALGNMYELKH 934

Qy   455 DNPSLKSPPPGFDSVIARGHTEPDTQDTELELDGQQVVVPQGQPVCPEFSSSTFSQSE 514
Db   935 AS-HISKLPKGKHSVKGLGKTPPDS--ANISLDG--VDVPLGTGI-SSGVNDTSLLYNE 988

Qy   515 YLIYQESQCRLRYLLEV 531
Db   989 YIVYDIAQVNLKYLLKL 1005


RESULT 8
S26057
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 18-Jun-1999
A;Cross-references: S26057; S78453; I52331; I52331
R;Thibodeau, J.; Gradwohl, G.; Dumas, C.; Clairoux-Moreau, S.; Brunet, G.; Penning, C.;
Biochem. Cell Biol. 67, 653-660, 1989
A;Title: Cloning of rodent cDNA encoding the poly(ADP-ribose) polymerase catalytic domai
A;Reference number: I52331; MUID:90027702
A;Accession: S26057
A;Molecule type: mRNA
A;Residues: 1-500 <THI>
A;Cross-references: EMBL:X65497; NID:g56849; PIDN:CAA46478.1; PID:g56850
R;Potvin, F.
submitted to the EMBL Data Library, March 1992
```

```
A;Reference number: S78453
A;Accession: S78453
A;Molecule type: mRNA
A;Residues: 1-124,'H',126-127,'A',129-238,'D',240-500 <POT>
A;Cross-references: EMBL:X65497
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltrans

Query Match    23.3%;  Score 657;  DB 2;  Length 500;
Best Local Similarity    34.4%;  Pred. No. 2.6e-38;
Matches 177;  Conservative 96;  Mismatches 190;  Indels 52;  Gaps 21;

Qy    41 AEKRI------IRVDPTCPLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQ 89
Db     5 SEKRMKLTLKGGAAVDPDSGLEHS-AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLE 62

Qy    90 -DSNRFFTCWNRWGRVGEV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHP 147
Db    63 SDKESRYWIFRSWGRVGTVIGSNKLEQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTKYP 121

Qy   148 GKYTLIEVQ-AEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTM 206
Db   122 KKFYPLEIDYGQDE--EAVKKL---AVKPGTK----SKLPKPVQELVGMIFDVESMKKAL 172

Qy   207 ALMDLDVKKMPLGKLSKQQIARGFEALEALEAALKGPTDGGQSLEELSSHFYTVIPHNFG 266
Db   173 VEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQAVSQGSSESQIL-DLSNRFYTLIPHDFG 231

Qy   267 HSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQL 326
Db   232 MKKPPLLNNTDSVQAKVEMLDNLLDIEVAYSLLRGGSDDSSKD---PIDVNYEKLKTDI 287

Qy   327 QLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ-HIWKVNQEGEDRFQAHSKLGNRKLLW 383
Db   288 KVVDRDSEEAEVIRKYVKNTHATTHNAYDLEVIDFKIEREGESQRYKPFRQLHNRRLLW 347

Qy   384 HGTNMAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---V 436
Db   348 HGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANY----CHTSQGDPI 402

Qy   437 GYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDTQDTELELDGQQVVVPQ 496
Db   403 GLILLGEVALGNMYELKHAS-HISKLPKGKHSVKGLGKTAPDDS--ASITLDG--VEVPL 457

Qy   497 GQPVCPEFSSSTFSQSEYLIYQESQCRLRYLLEV 531
Db   458 GTGIP-SGVNDICLLYNEYIVYDIAQVNLKYLLKL 491


RESULT 9
S31735
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - African clawed frog (fragment)
N;Alternate names: poly ADP-ribose polymerase
C;Species: Xenopus laevis (African clawed frog)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S31735; PN0495
R;Saulier-le Drean, B.M.
submitted to the EMBL Data Library, May 1992
A;Reference number: S31735
A;Accession: S31735
A;Molecule type: mRNA
A;Residues: 1-998 <SAU>
A;Cross-references: EMBL:Z12139; NID:g64967; PIDN:CAA78126.1; PID:g1334661
R;Ozawa, Y.; Uchida, K.; Okada, M.; Aml, Y.; Kushida, S.; Okada, N.; Miwa, M.
Biochem. Biophys. Res. Commun. 193, 119-125, 1993
A;Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)polymera
A;Reference number: PN0494; MUID:93277538
A;Accession: PN0495
A;Molecule type: mRNA
A;Residues: 742-745,'E',747-876 <OZA>
C;Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and diff
C;Superfamily: NAD+ ADP-ribosyltransferase
```

```
C;Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pentosyl

Query Match            23.1%;    Score 653;  DB 2;   Length 998;
Best Local Similarity  32.5%;    Pred. No. 1.3e-37;
Matches 177;  Conservative 102;  Mismatches 214;  Indels 52;  Gaps 20;

Qy      3 PKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQV 62
            :||:| ||||: : ||| : ||:::||::    ::|:|   || ::::::::::
Db    480 PSSGPVAGKSSGKVKEEKGSNKSEKKMKLTVKGGAIDPDSEL--EDSCHVLETGG--- 533

Qy     63 YEDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEV-GQSKINHFTRLED 120
            |:| ||||||||||||||||| | |  :|||||:||||||||| | | |||||||||
Db    534 -KIFSATLGVDITRGTNSYYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKED 592

Qy    121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 180
            ||| || ||: :|| ||| ||: :| ||||:|:||  ::| || |||:|||| ||| |:| :|:|
Db    593 AIEHFLNLYQDKTGNAW-HSPNFTKYPKKFYPLEIDYGQE-EDVVKKLSVG-AGTKSKLA 649

Qy    181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240
            |  || || :|||: || :: ||| : |  ||:  || |||:|||| ::| :|:|| ||
Db    650 KP------VQELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAV 703

Qy    241 KGPTDGGQSLEE-----LSSHFYTVIPHNFGHSOPPPINSPELLQAKKDMLLVLADIELA 295
            || ||||:|:|      |:||||:||| :||:|| |||:||||||||||:| ||||:|:
Db    704 ------SESLSEARLLDLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVA 757

Qy    296 QAL--QAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRC 352
            ||   |:|:|: || ::|:::|  :| || :::|||||||:|  || || ||  :|||||
Db    758 YSLLRGGADDGEK------DPIDVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNA 811

Qy    353 PTLQ--HIWKVNQGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SG 406
            ||    |||||||||||| ||:|||||||||:||||| |||||:||:||||        |
Db    812 YDLEVLEIFKIDREGEYQRYKPFKQLHNRQLLWHGSRTNFAGILSQGLRIAPPEAPVTG 871

Qy    407 GRVGKGIYFASENSKSAGVIGMKCGAAHVGYMFLGEVALGREHHINTDNPSLKSPPPGF 466
            |||||||||||||||||||||||| |  |||:|||||||||:|||||||||||||   |
Db    872 YMFGKGIYFADMVSKSANYCHAMP--GSPIGLILLGEVALGNMIELKAASQITKL-PKGK 928

Qy    467 DSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLR 526
            |:|||||:|||||| |||||:|||| :||:|::  ||:||  : ::|::||| ::  |||
Db    929 HSVKGLGRTAPDPS--ATVQLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLK 983

Qy    527 YLLEV 531
          ||:::
Db    984 YLLKL 988
```

```
RESULT 10
S04200
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S04200
R;Huppi, K.; Bhatia, K.; Siwarski, D.; Klinman, D.; Cherney, B.; Smulson, M.
Nucleic Acids Res. 17, 3387-3401, 1989
A;Title: Sequence and organization of the mouse poly (ADP-ribose) polymerase gene.
A;Reference number: S04200; MUID:89263780
A;Accession: S04200
A;Molecule type: DNA
A;Residues: 1-1013 <HUP>
A;Cross-references: EMBL:X14206; NID:g49893; PIDN:CAA32421.1; PID:g49894
C;Genetics:
A;Map position: 1
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc fi

Query Match            23.0%;    Score 649.5; DB 1;   Length 1013;
Best Local Similarity  32.4%;    Pred. No. 2.4e-37;
Matches 180;  Conservative 103;  Mismatches 205;  Indels 67;  Gaps 22;

Qy      1 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRI------IRVDPTCP 53
```

Second column (right side):

```
           :||:| ||                  :||:|                          |||
Db    493 VAPRGK-----------SAAPSKKSKGCFKEEGVNKS------------EKRMKLTLKGGAAVDPDSG 537

Qy     54 LSSNPGTQYYED---YNCTLNQTNIENNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV- 107
            |:||: | :||    ||||||||||||||||||||||| |||| | ||||||||||
Db    538 LEHS- AHVLEKGGKVFSATIGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRLGTVI 595

Qy    108 GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVV 166
            ||:||||||||||:|||||||||||||:||||||||||||:::|:||:||   ||:|  |
Db    596 GSNKLEQMPSKEEAVEQFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVK 652

Qy    167 KVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQI 226
            | ||| ||| ||:    ::|: |||||||||:|||||||||||||||||||||||:|||
Db    653 KLTVKP-GTKSKLPKP------VQELVGMIFDVDSMKKALVEYEIDLQKMPLGKLSRRQI 705

Qy    227 ARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDML 286
            |: |||||||:|||: ||:|:|: | ::::|||| :|| | |  || ||||| | :|:|:
Db    706 QAAYSILSEVQQPVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEML 764

Qy    287 LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 346
            | :||:|::: ||:|:||| ::::| ||  :| | ::: :|:::|::|||   | :| :|
Db    765 DNLLDIEVAYSLLRGGSDDSSKD--PIDVNYEKLKTDIKVVDRDSEEAEVIRKTVKNT 820

Qy    347 -GSNHRCPTLQ--HIWKVNQGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP 403
             :::||  ||    :||||||||:||:|||:|||::|||||:|  |||||::|||||
Db    821 HATTHNAYDLEVIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTNFAGILSQGLRIAP 880

Qy    404 H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDN 456
             |    | ||||||||:||:|| :||| :|| : ||    |||:||||||||:|||
Db    881 PEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLIMLGEVALGNMYELKHAS 935

Qy    457 PSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYL 516
            |:|:| ::|  | |||  ||| |  ||||  :|:|| ::||:|: |  ||::  |  :|:
Db    936 -HISKLPKGKHSVKGLGKTTPDPSASITLE---GVEVPLGTGIP-SGVNDTALLYNEYI 989

Qy    517 IYQESQCRLRYLLEV 531
            || :| |||:|||:
Db    990 VYDIAQVNLKYLLKL 1004
```

```
RESULT 11
A47474
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: A47474
R;Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.
Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993
A;Title: Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase: leucine zip
A;Reference number: A47474; MUID:93234521
A;Accession: A47474
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-994 <UCH>
A;Cross-references: GB:D13806; GB:D13807; GB:D13808; NID:g303545; PIDN:BAA02964.1; PI
A;Note: sequence extracted from NCBI backbone (NCBIN:129703, NCBIP:129704)
C;Genetics:
A;Gene: FlyBase:Parp
A;Cross-references: FlyBase:FBgn0010247
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match            22.68%;   Score 638;  DB 1;   Length 994;
Best Local Similarity  32.2%;   Pred. No. 1.5e-36;
Matches 168;  Conservative 100;  Mismatches 195;  Indels 58;  Gaps 19;

Qy     37 KAIPAEKRI-----IRVDPTCPLSSNPGTQVYED----YNCTLNQTNIENNNKFYIIQL 87
            | |||||:      ||||| |:|||||| ::|:||     :|||:|: :||:|
Db    500 KSMPVSRTFKVKDGLAVDPDSGLEDI-AHYVDSNNKYSVVLGLTDIQRNKNSYKVQL 557

Qy     88 LQ-DSNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVS 145
            ||  |||||:|||||||   | |: ||||  |::||||| :||| ::|:
```

```
Db  558  LKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVK  617

Qy  146  HPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNT  205
              ::  :||| ||   ::|:|
Db  618  RTGRMYPIEIQYDDD--QKLVKHE-------SHFFTSKLEISVQNLIKLIFDIDSMNKT   667

Qy  206  MALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNF  265
         : :|| |:|||||||| : ||   :  :  :||    ||| : :||::|||||:||||||
Db  668  LMEFHIDMDKMPLGKLSAHQIQSAYRVKEIYNVLECGSNTAK-LIDATNRFYTLIPHNF   726

Qy  266  GHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQ  325
         | | | :  :||: |    ||  ||||:  :|  | ::  |  |::    |||:|  :  |
Db  727  GVQLPTLIETHQOIEDLRQMLDSLAEIEVAYSI----IKSEDVSDACNPLDNHYAQIKTQ  782

Qy  326  LQLLDSSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLL  382
         |  ||:: | ||:||::|::   ||:    :   ||::|: |  |  :|  | |||:|||
Db  783  LVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARRFKPFKKLHNRKLL   842

Qy  383  WHGTNMAVVAAILTSGLRI----MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY  438
         |||:: :: || ||:|||     ||||||| ||||||:::|||||:|  :::       |
Db  843  WHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFADMVSKSANYCCTSQ--QNSTGL  900

Qy  439  MFLGEVALG-------REHHIN--TDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQ  490
         | |:|||||       :|||    ||||  |       ||  | | |||: : ::   ||
Db  901  MLLSEVALGDMMECTSAKYINKLSNNK--------HSCFGRGRTMPDPTK-SYIRSDG-   949

Qy  491  QVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEV  531
             :  |  :   ::| :  : :|:||  :    ::::|:|
Db  950  -VEIPYGETI-TDEHLKSSLLYNEYIVYDVAQVNIQYLFRM  988


RESULT 12
T51353

NAD+ ADP-ribosyltransferase (EC 2.4.2.30) [imported] - Arabidopsis thaliana
N;Alternate names: poly(ADP-ribose) polymerase
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C;Accession: T51353
R;Doucet-chabeaud, G.; Kazmaier, M.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z25379
A;Accession: T51353
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-983 <DOU>
A;Cross-references: EMBL:AJ131705; PIDN:CAA10482.1
A;Experimental source: cultivar landsberg erecta
C;Genetics:
A;Gene: parp-1
C;Function:
A;Description: ADP-ribose polymer synthesis
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match   21.2%;   Score 598.5;   DB 2;   Length 983;
Best Local Similarity   31.1%;   Pred. No. 8.7e-34;
Matches  171;   Conservative 101;   Mismatches 199;   Indels 79;   Gaps 23;

Qy   17  KKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQ----VYED----YNC   68
         || ||    : || |  :::     |:  | :      | :::||        |    ||
Db  472  KKQRKLPFDKYKIEDTSESLVTVKVKGR------SAVHEASGLQEHCHILEDGNSIYNT   524

Qy   69  TLNQTNIENNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD  124
         || ::::  :|:::|| | ::  ::  |   :||||||  | |::|:        |::|
Db  525  TLSMSDLSTGINSYIILQIIQE-DKGSDCYVFRKWGRVGNEKIGGNKVEEMSK-SDAVHE  582

Qy  125  FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ  181
         |::  :|||:|   ||   |    : || : :||  | || |:: |           :
Db  583  FKRLFLEKTGNTWESWEQKTNFQKQPGKFLPLD---------IDYGVNKQVAKK-E     628

Qy  182  P----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE  237
         |    |::  |: : | ::: |||:   || : |||||||||||||::|| |||||  :
```

```
Db  629  PFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ  688

Qy  238  EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIE   293
         ||     ||||   |||||||| :|:|:|:|| | : :  | :  |:|:|:|:|:||||
Db  689  RLLTESDPQPTMKESLLVDASNRFFTMIP----SIHPHIIRDEDDFKSVKMLEALQDIE   744

Qy  294  LAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP  353
         :|    |:  :  | :      : |:   | | : :|::   : ::  || ::   :
Db  745  IASRI--VGFDVDSTES----LDDKYKKKLHCDISPLPHDSEDYRLIEKYLNTT---HAP  794

Qy  354  T------LQHIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRI----  401
                :::|:|||:|  | |:: ||  :|||::::|||| :: ::|   | |
Db  795  THTEWSLEEEVFALERGEGFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPE  854

Qy  402  MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS  461
          | :| :::||||||:: :||:|||  :::  ||||||::| :|| || | :  || |
Db  855  APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGETHEL-TKAKYMDK  911

Qy  462  PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES  521
         || | :     |  :|:    : |  |    |||  : |||   :  ::|::|::||: :
Db  912  PPRGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEYIVYDTA  967

Qy  522  QCRLRYLLEV  531
              :::||:|
Db  968  QVKLQFLLKV  977


RESULT 13
C84719

probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: C84719
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: C84719
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1009 <STO>
A;Cross-references: GB:AE002093; NID:g4432827; PIDN:AAD20677.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g31320
A;Map position: 2
C;Superfamily: NAD+ ADP-ribosyltransferase

Query Match   21.2%;   Score 598.5;   DB 2;   Length 1009;
Best Local Similarity   31.1%;   Pred. No. 9e-34;
Matches  171;   Conservative 101;   Mismatches 199;   Indels 79;   Gaps 23;

Qy   17  KKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQ----VYED----YNC   68
         || ||    : || |  :::     |:  | :      | :::||        |    ||
Db  498  KKQRKLPFDKYKIEDTSESLVTVKVKGR------SAVHEASGLQEHCHILEDGNSIYNT   550

Qy   69  TLNQTNIENNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD  124
         || ::::  :|:::|| | ::  ::  |   :||||||  | |::|:        |::|
Db  551  TLSMSDLSTGINSYIILQIIQE-DKGSDCYVFRKWGRVGNEKIGGNKVEEMSK-SDAVHE  608

Qy  125  FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ  181
         |::  :|||:|   ||   |    : || : :||  | || |:: |           :
Db  609  FKRLFLEKTGNTWESWEQKTNFQKQPGKFLPLD---------IDYGVNKQVAKK-E     654

Qy  182  P----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE  237
         |    |::  |: : | ::: |||:   || : |||||||||||||::|| |||||  :
Db  655  PFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ  714

Qy  238  EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIE   293
         ||     ||||   |||||||| :|:|:|:|| | : :  | :  |:|:|:|:|:||||
Db  715  RLLTESDPQPTMKESLLVDASNRFFTMIP----SIHPHIIRDEDDFKSVKMLEALQDIE   770
```

```
Db  791 EVFSLDRDGELNKYSRYKNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGK 850

Qy  412 GIYFASENSKSAGYVIGMKCGA---HHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDS 468
Db  851 GLYFADLVSKSAQY-----CYVDRNNPVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS 904

Qy  469 VIARGHTEPDTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYL 528
Db  905 TKGLGKTVP--LESEFVKWRDDVVVPCGKPVP-SSIRSSELMYNEYIVNTSQVKMQFL 960

Qy  529 LEV 531
Db  961 LKV 963


RESULT  15
T18600
hypothetical protein AC8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18600
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z18996
A;Accession: T18600
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-727 <WIL>
A;Cross-references: EMBL:Z83097; PIDN:CAB05448.1; GSPDB:GN00023; CESP:AC8.1
A;Experimental source: clone AC8
C;Genetics:
A;Gene: CESP:AC8.1
A;Map position: 5
A;Introns: 31/3; 271/3; 312/3; 348/1; 466/3; 545/3; 617/3; 702/2


Query Match     18.9%;    Score 534;  DB 2;   Length 727;
Best Local Similarity   28.6%;    Pred. No. 1.9e-29;
Matches 162;   Conservative 104;   Mismatches 211;   Indels  90;   Gaps   20;

Qy   30 RSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQ 89
Db  182 KATGEYVEALAAKGG--STEPATPASASP-TPPEAETPVLSAEGSPESSNKRPASAEIIE 238

Qy   90 ------DSNRFF-TCWNRWGRVGEVGQSKI------------------NHFTRLED 120
Db  239 IDGEGNPDENDFAKKRRMKKEARLMEVQKKRMKKQSDLLWEYRQIFERMPYTDNISILRE 298

Qy  121 AKKDF--------EKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGP 172
Db  299 NEQDIPEGHDRTAQDVFHEKTKNDWIYRKHFRKMPGMFSYVETDYSE----FVGTNNGH 353

Qy  173 VRTVT--KRVQPCS---LDPATQKLITNIFSKEMPKNTMALMDLDVKKMPLGKLSKQQIA 227
Db  354 KKKITPGSKITPGSKTLLPFKSVKEVMSIFDVENMKSALKSFEIDVNKMPLGRLSHNQIN 413

Qy  228 RGFEALEALEAL-KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDML 286
Db  414 LAFEVLNDISDLLVKLPIDASKIL-DFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNML 472

Qy  287 LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 346
Db  473 NALLDIKFAYDQISGGDVPASTLSLSIDPVDINYRKLKCIMEPLQQGCDDWNMIHQYLKNT 532

Qy  347 -GSNHRCPT-LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH 404
Db  533 HGATHDLKVELIDILKVNRDNESSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPP 590

Qy  405 ---SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGR-EHHINTDNPSL 459
Db  591 EAPVSGYMFGKGVYFADMFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSR 647

Qy  460 KSPPPGFDSVIARG--------PTQDTELELDGQQVVVPQGQPVPCPEFS 506
```

```
Qy  294 LAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP 353
Db  771 IASRI--VGFDVDSTES----LDDKYKKLHCDISPLPHDSEDYRLIEKYLNTT----HAP 820

Qy  354 T------LQHIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRI---- 401
Db  821 THTEWSLELEEVFALERERGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPE 880

Qy  402 MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS 461
Db  881 APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGEIHEL-TKAKYMDK 937

Qy  462 PPGFDSVIARGHTEPDTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES 521
Db  938 PPRGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEYIVDTA 993

Qy  522 QCRLRYLLEV 531
Db  994 QVKLQFLLKV 1003


RESULT  14
T03657
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) 2 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
C;Accession: T03657
R;Babiychuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Monta
submitted to the EMBL Data Library, November 1997
A;Reference number: Z14992
A;Accession: T03657
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-969 <BAB>
A;Cross-references: EMBL:AJ222589; PIDN:CAA10889.1
C;Genetics:
A;Gene: PARP2
C;Superfamily: NAD+ ADP-ribosyltransferase
C;Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase


Query Match     20.5%;    Score 578;  DB 2;   Length 969;
Best Local Similarity   30.6%;    Pred. No. 2.3e-32;
Matches 166;   Conservative 96;   Mismatches 195;   Indels  86;   Gaps   21;

Qy   16 KKKGRQAGREEDPFPRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNI 75
Db  480 KVKGRSAVHESSGLQDTAHILE----------DGKSI---YNATLNMSDL 516

Qy   76 ENNNNKFYIIQLL-QDSNRFFTCWNRWGRVG--EVGQSKINHFTRLEDAKKDFEKKFREK 132
Db  517 ALGVNSYYVLQIIEQDDGSECYVPRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEK 575

Qy  133 TKNNWAE--RDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPAT 189
Db  576 TGNSWEAWECKTNFRKQPGRFYPLDVD--------YGVKKAPKRKDISEMK-SSLAPQL 625

Qy  190 QKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEALKGPTDGGQS 249
Db  626 LELMKMLFNVETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALA 685

Qy  250 LEE------LSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAV-SE 303
Db  686 VRESLIVAASNRFFLIP---SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSD 741

Qy  304 QEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT------LQ 356
Db  742 SDES------LDDKYMKLHCDITPLAHDSEDYKLIEQYL--LNTHAPTHKDWSLELE 790

Qy  357 HIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGK 411
```

```
Db    648 QTLPAGFQSVQGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQGKQNV------------ 694
           ::.|.||.||  |          :      ||       |.::| |:| |
Qy    507 SSTFSQSEYLIYQESQCRLRYLLEVHL 533
           :|:::|  |.:||:|.| :
Db    695 DYHLLYNEFIVYDVDQIQLKYLVRVKM 721
```

Search completed: August 29, 2002, 07:59:29
Job time: 276 sec

```
                    GenCore version 4.5
            Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      August 29, 2002, 07:58:18 ; Search time 46.79 Seconds
                                          (without alignments)
                                          278.240 Million cell updates/sec

Title:       US-09-701-586B-8
Perfect score: 2813
Sequence:    1 MAPKKRASVQTEGSKKQRQG.........EYLIYKESQCRLRYLLEIHL 533

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:    231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters:      231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
             1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
             2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
             3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
             4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
             5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
             6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                           SUMMARIES

              %
Result       Query
 NO.   Score Match Length DB  ID                  Description
-----  ----- ----- ------ --  --------------       -----------
  1     705   25.1   1013   4  US-08-860-886-2      Sequence 2, Appli
  2     700   24.9   1014   4  US-09-078-347A-3     Sequence 3, Appli
  3     146    5.2   1327   4  US-09-196-387-2      Sequence 2, Appli
  4     105.5   3.8   635   4  US-09-046-992-2      Sequence 2, Appli
  5     104     3.7  2482   1  US-08-328-254-6      Sequence 6, Appli
  6     103.5   3.7   613   1  US-08-405-615-1      Sequence 1, Appli
  7     103.5   3.7   613   2  US-08-461-234-1      Sequence 1, Appli
  8     103.5   3.7   613   2  US-08-463-480-1      Sequence 1, Appli
  9     103.5   3.7   614   1  US-08-225-224-1      Sequence 1, Appli
 10     103.5   3.7   614   3  US-08-722-258-1      Sequence 1, Appli
 11     103.5   3.7   614   3  PCT-US95-04468-1     Sequence 1, Appli
 12     103.5   3.7   638   3  US-09-047-148-2      Sequence 2, Appli
 13     100.5   3.6   420   1  US-08-391-259-7      Sequence 7, Appli
 14     100.5   3.6   420   4  US-08-839-425-7      Sequence 15, Appl
 15     100.5   3.6   889   4  US-09-336-447A-15    Sequence 42, Appli
 16      99.5   3.5   462   2  US-08-484-438-42     Sequence 4, Appli
 17      99     3.5   497   3  US-08-511-485-4      Sequence 4, Appli
 18      99     3.5   497   3  US-09-212-971-4      Sequence 4, Appli
 19      99     3.5   497   4  US-08-800-929A-4     Sequence 16, Appl
 20      99     3.5   497   4  US-09-617-053A-4     Sequence 1, Appli
 21      99     3.5   622   2  US-08-356-786-16     Sequence 31, Appli
 22      99     3.5  3248   1  US-08-353-700-1      Sequence 1, Appli
 23      98.5   3.5  3248   5  PCT-US95-16216-1     Sequence 31, Appl
 24      98.5   3.5   579   5  PCT-US94-09143-31    Sequence 5, Appli
 25      98.5   3.5   579   5  US-08-460-309-5      Sequence 5, Appli
 26      98     3.5  3075   2  US-08-125-077-5      Sequence 5, Appli
 27      98     3.5  3075   2                       
```

```
 28   97.5   3.5   508   1  US-08-021-601-10     Sequence 10, Appl
 29   97.5   3.5   508   1  US-08-082-849B-10    Sequence 10, Appl
 30   97.5   3.5   508   5  PCT-US94-01624-10    Sequence 10, Appl
 31   97.5   3.5  2285   5  US-09-308-375-2      Sequence 2, Appli
 32          3.4   163   1  US-08-044-618-6      Sequence 6, Appli
 33          3.4   637   1  US-08-235-838-14     Sequence 14, Appl
 34          3.4   637   4  US-08-465-473B-14    Sequence 14, Appl
 35          3.4  1128   3  US-08-923-992A-6     Sequence 6, Appli
 36          3.4  1164   4  US-08-923-992A-2     Sequence 2, Appli
 37          3.4   396   4  US-09-046-992-4      Sequence 4, Appli
 38          3.4   420   1  US-08-391-259-2      Sequence 2, Appli
 39          3.4   420   2  US-08-391-259-11     Sequence 11, Appl
 40          3.4   420   2  US-08-839-425-2      Sequence 2, Appli
 41          3.4   420   2  US-08-839-425-11     Sequence 11, Appl
 42          3.4   599   1  US-08-463-163-3      Sequence 3, Appli
 43          3.4   637   2  US-08-235-838-16     Sequence 16, Appl
 44          3.4   637   2  US-08-465-473B-16    Sequence 16, Appl
 45          3.4   816   2  US-08-533-306A-6     Sequence 6, Appli
```

```
                              ALIGNMENTS

RESULT  1
US-08-860-886-2
; Sequence 2, Application US/08860886
; Patent No. 6335009
; GENERAL INFORMATION:
;  APPLICANT:  Burkle, Alexander
;  APPLICANT:  zur Hausen, Harald
;  APPLICANT:  Jan-Heiner, Kupper
;  TITLE OF INVENTION:  VECTORS AND VIRUSES FOR USE
;  TITLE OF INVENTION:  IN GENE THERAPY
;  NUMBER OF SEQUENCES:  2
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Pennie & Edmonds, LLP
;  STREET:  1155 Avenue of the Americas
;  CITY:  New York
;  STATE:  NY
;  COUNTRY:  USA
;  ZIP:  10036-2811
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  Diskette
;  COMPUTER:  IBM Compatible
;  OPERATING SYSTEM:  Windows
;  SOFTWARE:  FastSEQ for Windows Version 2.0b
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/860,886
;  FILING DATE:  03-OCT-1997
;  CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER:
;  FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  Coruzzi, Laura A
;  REGISTRATION NUMBER:  30,742
;  REFERENCE/DOCKET NUMBER:  8484-0028-999
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  650-493-4935
;  TELEFAX:  650-493-5556
;  TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH:  1013 amino acids
;  TYPE:  amino acid
;  STRANDEDNESS:  single
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  protein
;  FRAGMENT TYPE:  internal
US-08-860-886-2

Query Match   25.1%;  Score 705;  DB 4;  Length 1013;
```

```
Best Local Similarity   33.8%;   Pred. No. 1.7e-60;
Matches 183;  Conservative 101;  Mismatches 218;  Indels 40;  Gaps 17;

Qy    1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
Db  492 VAPRGKSGA--ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 548

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
Db  549 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKED 608

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
Db  609 AIEHFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL---TVNPGTKSK 663

Qy  178 VVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
Db  664 LPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQ 717

Qy  238 AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTL 297
Db  718 AVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSL 776

Qy  298 QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQ----TGNSYRCP 353
Db  777 LRGGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATHNAYDL- 830

Qy  354 NLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRV 409
Db  831 EVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTNFAGILSQGLRIAPPEAPVTGYMF 890

Qy  410 GKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSV 469
Db  891 GKGIYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGKHSV 947

Qy  470 IARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLL 529
Db  948 KGLGKTPDPSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEYIVYDIAQVNLKYLL 1002

Qy  530 EI 531
Db 1003 KL 1004
```

```
Query Match           24.9%;   Score 700;  DB 4;  Length 1014;
Best Local Similarity 33.6%;   Pred. No. 5.3e-60;
Matches 182;  Conservative 102;  Mismatches 219;  Indels 38;  Gaps 17;

Qy    1 MAPKRKASVQTEGSKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQV 60
Db  493 VAPRGKSGA--ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLEKG 549

Qy   61 HEDYDCTLNQTNIGNNNNKFYIIQLLEEG--SRFFCWNRWGRVGEV-GQSKMNHFTCLED 117
Db  550 GKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKED 609

Qy  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKALSPQVDSGPVRT 177
Db  610 AIEQFMKLYEEKTGNAWHSKN-FTKYPKKFYPLEID-YGQDEEAVKKL--TVNPGTKSK 664

Qy  178 VVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEE 237
Db  665 LPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQ 718

Qy  238 AMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTL 297
Db  719 AVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSL 777

Qy  298 QAAPGEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT-GNSYRCPNLR 356
Db  778 LRGGSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATHSAYDLE 832

Qy  357 --HVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVG 410
Db  833 VIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTNFAGILSQGLRIAPPEAPVTGYMFG 892

Qy  411 KGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVI 470
Db  893 KGIYFADMVSKSANYYHTSQ--GDPIGLILLGEVALGNMYELK-HASHISRLPKGKHSVK 949

Qy  471 ARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLE 530
Db  950 GLGKTPDPSANI--SLDG--VDVPLGTGIS-SGVIDTSLLYNEYIVYDIAQVNLKYLLK 1004

Qy  531 I 531
Db 1005 L 1005
```

```
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,225
FILING DATE: June 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-196-387-2
```

```
Query Match      5.2%;   Score 146;  DB 4;  Length 1327;
Best Local Similarity  20.2%;  Pred. No. 4.2e-05;
Matches  123;  Conservative  83;  Mismatches 232;  Indels 170;  Gaps  27;

Qy    14 SKKQRQG----------TEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGIQVHE 62
Db   777 TKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAA-KKGCLARVQKLC-----------TPE 826

Qy    63 DYDCTLNQTNIGNNNNKFYII---------QLLEEGSRFFCWNRWNGRVG---------- 102
Db   827 NINCRDTQ--GRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHV 883

Qy   103 EVGQSKMNHFTCLEDAKKDFKKKFWEKT------KNKWEERDRFVAQPNKYTLIEVQGE 155
Db   884 DIAALLIKYNTCVNATDK------WAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQ 937

Qy   156 A--------ESQEAVVKALSPQ---VDSGVPRTVVKPCSLDPATQNLITNIFSKEMFKNA 204
Db   938 TPLDLATADDIRALLIDAMPEALPTCFKPQATVVSASLISPAS----------TPSCLSAA 989

Qy   205 MTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGDGQSLEELSSCFYTVIPHNF 264
Db   990 SSIDNL---TGPLAELAVGGASNAGDGAAGTE-----RKEGEVAGLDMNISQFLKSLGLEH 1042

Qy   265 GRSRPPPINSPDVLQAKKDMLLVLADI---ELAQTLQAAPGEEEEKVEEVPHPL----D 316
Db  1043 LR---------DIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGCQQGT 1093

Qy   317 RDYLLRCQLQ--LLDSG--ESEYKAIQTYLKQT--------GNSYRCPNLRHVWKV 361
Db  1094 NPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKV 1153

Qy   362 NREGEGDRF----QAHSKLGNRRLLWHGTNVAVVAAILTSGL-RIMPHSGGRVGKG 412
Db  1154 VNKKLRERFCHRQKEVSSEENHNHNERMLFHGS--PFINAIIHKGFDERHAYIGGMFGAG 1211

Qy   413 IYFASENSKSAGYVTTM------HCGGHQ------VGYMFLGEVALGKEHHITIDDPSLKS 461
```

```
Db  1212 IYFAENSSKSNQVYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKS-FLQFSTMKMAH 1270
Qy   462 PPPGFDSVIARGTEPDPAQDIELELDGGPVVPQGPPVQCPSFKSSSFSQSEYLIYKES 521
Db  1271 APPGHHSVI--------GRPSV--------------------------NGLAYAEYVIYRGE 1298

Qy   522 QCRLRYLL 529
Db  1299 QAYPEYLI 1306
```

```
RESULT 4
US-09-046-992-2
; Sequence 2, Application US/09046992
; Patent No. 6140066
; GENERAL INFORMATION:
; APPLICANT: Lorberboum-Galski, Haya
; APPLICANT: Yarkoni, Shai
; APPLICANT: Ben-Yehudah, Ahmi
; TITLE OF INVENTION: METHODS OF CANCER DIAGNOSIS
; TITLE OF INVENTION: USING A CHIMERIC TOXIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,992
; FILING DATE: 24-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 9457-0013-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-5556
; TELEFAX: 650-493-4935
; TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-046-992-2
```

```
Query Match      3.8%;   Score 105.5;  DB 4;  Length 635;
Best Local Similarity  19.7%;  Pred. No. 0.13;
Matches  78;  Conservative  48;  Mismatches 143;  Indels 127;  Gaps  17;

Qy   172 SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 213
Db   214 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISE 268

Qy   214 ------------------------KMPLGKLTKQQIARGFEALE-------- 233
Db   269 ELEFPEGGSLAALTAHQACHLPLETFTHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 328

Qy   234 ----ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 289
```

```
Db  329 QVDQVIRNALASP-GSGGDIGEAI---------REQPE--------QAR--LALTLA 365
          :|: :        | ||:                 |:|             :||
Qy  290 DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAI 339
          ||: | ||:  |:                      |||       |||    |:|
Db  366 AAESRFVRQGTGNDEAGA-----ANADVVSLTCPVAAGECAGPADSGDALLEANYPTG 419
Qy  340 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL---WHGTNVAVVAAILT 396
          |||:|                 |:|| :|                 ::|  | :|
Db  420 AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGVYHGTFLEAAQSIVF 474
Qy  397 SGLRIMPHSGGRVGKGIYFASENSKSAGYTTMHCGGHQVGYMFLGEVALGKEHITIDD 456
           :|               :||:::: |: :      |||    |:
Db  475 GGVRARSQDLDAIWRGFYIAGDPALAYGIYAQDQEPDAR------GRIRNGALLRVYVPR 527
Qy  457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 492
          ||       |    :|   |   :|:||
Db  528 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 557
```

RESULT 5
US-08-328-254-6
; Sequence 6, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
;   APPLICANT: Zhu, Xueliang
;   APPLICANT: Lee, Wen-Hwa
;   TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Campbell and Flores
;   STREET: 4370 La Jolla Village Drive, Suite 700
;   CITY: San Diego
;   STATE: California
;   COUNTRY: USA
;   ZIP: 92122
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/328,254
;   FILING DATE: 24-OCT-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/141,239
;   FILING DATE: 22-OCT-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Campbell, Cathryn A.
;   REGISTRATION NUMBER: 31,815
;   REFERENCE/DOCKET NUMBER: P-CJ 1191
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 535-9001
;   TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2482 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-328-254-6

```
Query Match     3.7%;     Score 104; DB 1; Length 2482;
Best Local Similarity 21.4%;  Pred. No. 1.6;
Matches   80; Conservative 65; Mismatches 139; Indels 90; Gaps 17;

Qy   60 VHEDYD-CTLNQTNIGN----NNNKFYIIQLLEEGSRFFCWNRWRGVGEVGQSKMNHFT 113
          :|:: :  :|: |::         :|| : :||                   :|
Db 1825 LHNDQEACKAKEQNLSSQVECLELEKAQLLQGLDEAKNNYI-------VLQSSVNGLI 1875
Qy  114 .C-LEDAKKDFKKKFWE--KTKNKWEERDRFVAQPNKYTLIEVQGEAE--SQEAVVKALS 167
```

```
Db  1876 QEVEDGKQKLEKKDEISRLKNQIQDQEQLVSK----LSQVEGEHQLWKEQNLERNLT 1930
           :||  |:| :  |: :: :    |:            |:||||:||||:|||:|:||
Qy   168 PQVDSGPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIAR 227
           :::  :|  | |:|::|               |:| :|  :|:  |:
Db  1931 VELEQIQVLQSKNASLQDTLEVLQS---SYKNLENELELTKMD-KMSFVEKVNKMTAK 1985
Qy   228 GFEALEALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDML-- 285
           ||  | |   :::|   |:                          :|  |||:
Db  1986 ETELQRMEHE-MAQKTAELQ-EELSG------EKNRLAGEIQLLLEEIKSSKDQLKE 2034
Qy   286 LVLADIELAQTLQAAPGEEEKVEEVPHPLDRDYQL---LRCQLQLLDSGESEYKAI 339
           |: |:|: |:|| |      :|        |:|      |||||| |:|:|
Db  2035 LTLENSELKKSLDCMHKDQVEKEGKVREEI-AEYQLRLHEAEKKHQALLLDTNKQYEVEI 2093
Qy   340 QTYLKQTGNSYRC---------------------------------PNLRHVWKVN 362
           ||| ::|                                        |:| ||:
Db  2094 QTYREKLTSKEECLSSQKLEIDLLKSSKEELNNSLKATTQILEELKKTKMDNLKYVNQLK 2153
Qy   363 REGEGDRFQAHSKL 376
           :|    || |:|||
Db  2154 KENE--RAQGKMKL 2165
```

RESULT 6
US-08-405-615-1
; Sequence 1, Application US/08405615
; Patent No. 5602095
; GENERAL INFORMATION:
;   APPLICANT: Pastan, Ira
;   APPLICANT: FitzGerald, David J.
;   TITLE OF INVENTION: Recombinant Pseudomonas Exotoxin with
;   TITLE OF INVENTION: Increased Activity
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Ellen L. Weber
;   STREET: One Market Plaza, Steuart Tower, Suite 2000
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94105
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/405,615
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/07/901,709
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Weber, Ellen L.
;   REGISTRATION NUMBER: 32,762
;   REFERENCE/DOCKET NUMBER: 15280-36
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-543-9600
;   TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 613 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-405-615-1

```
Query Match     3.7%;     Score 103.5; DB 1; Length 613;
Best Local Similarity 19.7%;  Pred. No. 0.19;
Matches   78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;
```

STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-461-234-1

Query Match   3.7%;   Score 103.5;   DB 2;   Length 613;
Best Local Similarity 19.7%; Pred. No. 0.19;
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;

```
Qy   172 SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 213
             ||  |  :|::|                :|:||           :||:|
Db   192 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 246

Qy   214 -----------------------------KMPLGKLTKQQIARGFEALE---------- 233
                                          ||| |::| || :|:|  ||
Db   247 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 306

Qy   234 ---ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 289
            ||  |  ::| | || |        |       :|          |  ||: : ||
Db   307 QVDQVIRNALASP-GSGGDLGEAI-----------------REQPE------QAR--LALTLA 343

Qy   290 DIELAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQL----------LDSG----ESEYKAI 339
            :|:||                    |   || :||  |||          |:|      ::|
Db   344 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 397

Qy   340 QTYLKQTGNSYRCPNLRHVHWKVNREGEGDRFQAHSKLGNRRLL----WHGTNVAVVAAILT 396
            :|:    |     :||:|       :|||   |    |||:|||          |:|  ||:|
Db   398 AEFLGDGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 452

Qy   397 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 456
             |:|               || ||::||       :    |       :||: :|:
Db   453 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 505

Qy   457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 492
             || :| |||     |  :| | :||:|:
Db   506 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 535
```

RESULT 8
US-08-463-480-1
; Sequence 1, Application US/08463480
; Patent No. 5854044
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira H.
; APPLICANT: FitzGerald, David J.
; TITLE OF INVENTION: Recombinant Pseudomonas Exotoxin with
; TITLE OF INVENTION: Increased Activity
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,480
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,709
; FILING DATE: 18-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,615
; FILING DATE: 15-MAR-1995
; ATTORNEY/AGENT INFORMATION:

```
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 15280-36-2
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (415) 576-0200
  TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 613 amino acids
    TYPE: amino acid
    STRANDEDNESS:
    TOPOLOGY: linear
  MOLECULE TYPE: protein
  HYPOTHETICAL: NO
US-08-463-480-1

Query Match           3.7%;   Score 103.5; DB 2; Length 613;
Best Local Similarity 19.7%;  Pred. No. 0.19;
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;

Qy  172 SGPVRTVVKP---------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 213
Db  192 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH----DLDIKPTVISH 246

Qy  214 ---------KMPLGKLTKQQIARGFEALE------ 233
Db  247 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 306

Qy  234 ---ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLA 289
Db  307 QVDQVIRNALASP-GSGGDLGEAI---------REQPE------QAR--LALTLA 343

Qy  290 DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL------LDSG----ESEYKAI 339
Db  344 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 397

Qy  340 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL---WHGTNVAVVAAILT 396
Db  398 AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 452

Qy  397 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDD 456
Db  453 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 505

Qy  457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 492
Db  506 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 535

RESULT 9
US-08-225-224-1
; Sequence 1, Application US/08225224
; Patent No. 5635599
GENERAL INFORMATION:
  APPLICANT: PASTAN, Ira
  APPLICANT: KREITMAN, Robert J.
  TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
  TITLE OF INVENTION: CIRCULARLY PERMUTED FUSION PROTEINS
  NUMBER OF SEQUENCES: 57
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Townsend and Townsend Khourie and Crew
    STREET: Steuart Street Tower, One Market Plaza
    CITY: San Francisco
    STATE: California
    COUNTRY: US
    ZIP: 94105-1493
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
```

```
  APPLICATION NUMBER: US/08/225,224
  FILING DATE: 8-APR-1994
  CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
  NAME: Weber, Ellen L.
  REGISTRATION NUMBER: 32,762
  REFERENCE/DOCKET NUMBER: 15280-193
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (415) 543-9600
  TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 614 amino acids
    TYPE: amino acid
    STRANDEDNESS: unknown
    TOPOLOGY: unknown
  MOLECULE TYPE: protein
  FEATURE:
    NAME/KEY: Protein
    LOCATION: 1..614
    OTHER INFORMATION: /label= native-PE
US-08-225-224-1

Query Match           3.7%;   Score 103.5; DB 1; Length 614;
Best Local Similarity 19.7%;  Pred. No. 0.19;
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;

Qy  172 SGPVRTVVKP---------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 213
Db  193 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH----DLDIKPTVISH 247

Qy  214 KMPLGKLTKQQIARGFEALE 233
Db  248 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 307

Qy  234 ---ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 289
Db  308 QVDQVIRNALASP-GSGGDLGEAI---------REQPE------QAR--LALTLA 344

Qy  290 DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL------LDSG----ESEYKAI 339
Db  345 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 398

Qy  340 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL----WHGTNVAVVAAILT 396
Db  399 AEFLGDGGDVSFSTRGTQNWTVER-----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 453

Qy  397 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDD 456
Db  454 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 506

Qy  457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 492
Db  507 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 536

RESULT 10
US-08-722-258-1
; Sequence 1, Application US/08722258
; Patent No. 6011002
GENERAL INFORMATION:
  APPLICANT: Pastan, Ira
  APPLICANT: Kreitman, Robert J.
  APPLICANT: Puri, Raj K.
  TITLE OF INVENTION: Circularly Permuted Ligands and
  TITLE OF INVENTION: Circularly Permuted Chimeric Molecules
  NUMBER OF SEQUENCES: 72
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Townsend and Townsend and Crew LLP
    STREET: Two Embarcadero Center, Eighth Floor
    CITY: San Francisco
    STATE: California
```

COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,258
FILING DATE: 08-JAN-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/04468
FILING DATE: 06-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/225,224
FILING DATE: 08-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 015280-19310OUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 614 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..614
OTHER INFORMATION: /note= "native Pseudomonas exotoxin
OTHER INFORMATION:
US-08-722-258-1

Query Match   3.7%;   Score 103.5; DB 3;   Length 614;
Best Local Similarity 19.7%; Pred. No. 0.19;
Matches   78; Conservative   47; Mismatches 144; Indels 127; Gaps   17;

Qy   172 SGPVRTVVKP------------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 213
         ||:| :||: : || |: ||:|
Db   193 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 247

Qy   214 --------------------KMPLGKLTKQQIARGFEALE--- 233
                              ||:| |: ||:| ||
Db   248 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 307

Qy   234 ----ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 289
             || : :| ||| | || :|
Db   308 QVDQVIRNALASP-GSGGDLGEAI----------REQPE-------QAR--LALTLA 344

Qy   290 DIELAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAI 339
         ||| | ||| |
Db   345 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 398

Qy   340 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL---WHGTNVAVVAAILT 396
         ::| | | | |||:|: :   :||
Db   399 AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 453

Qy   397 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 456
         |:| :: : || : :|| :
Db   454 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 506

Qy   457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 492
         |:| ||: :| |||:| |:
Db   507 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 536

RESULT 11

PCT-US95-04468-1
Sequence 1, Application PCT/US9504468
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
TITLE OF INVENTION: CIRCULARLY PERMUTED FUSION PROTEINS
NUMBER OF SEQUENCES: 59
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04468
FILING DATE: 07-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/225,224
FILING DATE: 08-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen L.
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 15280-193-1PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 614 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..614
OTHER INFORMATION: /label= native-PE
PCT-US95-04468-1

Query Match   3.7%;   Score 103.5; DB 5;   Length 614;
Best Local Similarity 19.7%; Pred. No. 0.19;
Matches   78; Conservative   47; Mismatches 144; Indels 127; Gaps   17;

Qy   172 SGPVRTVVKP------------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK------ 213
         ||:| :||: : || |: ||:|
Db   193 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 247

Qy   214 --------------------KMPLGKLTKQQIARGFEALE--- 233
                              ||:| |: ||:| ||
Db   248 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 307

Qy   234 ----ALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLA 289
             || : :| ||| | || :|
Db   308 QVDQVIRNALASP-GSGGDLGEAI----------REQPE-------QAR--LALTLA 344

Qy   290 DIELAQTLQAAPGEEEEKVEEVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAI 339
         ||| | ||| |
Db   345 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 398

Qy   340 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL---WHGTNVAVVAAILT 396
         ::| | | | |||:|: :   :||
Db   399 AEFLGDGGDVSFSTRGTQNWTVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 453

Qy   397 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDD 456
         |:| :: : || : :|| :
Db   454 GGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPR 506

Qy   457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 492
         |:| ||: :| |||:| |:
Db   507 SSL----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 536

RESULT 12
US-09-047-148-2
; Sequence 2, Application US/09047148
; Patent No. 6086900
; GENERAL INFORMATION:
; APPLICANT: Draper, Rockford
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR USING
; TITLE OF INVENTION: MEMBRANE-PENETRATING PROTEINS TO CARRY MATERIALS ACROSS
; TITLE OF INVENTION: CELL MEMBRANES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/047,148
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/042,056
; FILING DATE: 26-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSF:072
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-047-148-2

Query Match   3.7%; Score 103.5; DB 3; Length 638;
Best Local Similarity   19.7%; Pred. No. 0.2;
Matches   78; Conservative   47; Mismatches 144; Indels 127; Gaps   17;

Qy   172 SGPVRTVVKP----------CSLDPATQNLITNIFSKEMFKNAMTLMNLDVK-------- 213
             || | ::| : | :   :||: :|
Db   217 SGKVLCLLDPLDGVYNYLAQQRCNLDDTWEGKIYRVLAGNPAKH------DLDIKPTVISH 271

Qy   214 ----------KMPLGKLTKQQIARGFEALE---------------------------------- 233
                     ||: ::| ||:|
Db   272 RLHFPEGGSLAALTAHQACHLPLETFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWN 331

Qy   234 ----ALEEAMKNPTGDGQSLEEILSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLA 289
                 ||: : |||: :| || : | || || : | |
Db   332 QVDQVIRNALASP-GSGGDLGEAI-----------REQPE--------QAR--LALTLA 368

Qy   290 DIELAQTLQAAPGEEEKVEEVPHPLDRDYQLLRCQLQL------LDSG----ESEYKAI 339
             | : | || : | :| | : | | :| :| | :||:||
Db   369 AAESERFVRQGTGNDEAGA------ANADVVSLTCPVAAGECAGPADSGDALLERNYPTG 422

Qy   340 QTYLKQTGNSYRCPNLRHVWKVNREGEGDRFQAHSKLGNRRLL---WHGTNVAVVAAILT 396
             : :| || | ||| :|||:|| :||:| |:|| ::
Db   423 AEFLGDGGDVSFSTRGTQNWTVER-----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVF 477

Qy   397 SGLRIMPHSGGRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEBHHITIDD 456
             |:| :| :| || :|| || | | | | |:
Db   478 GGVRARSQDLDAIWRGFYIAGDPALAYGTAQDQEPDAR------GRIRNGALLRVVVPR 530

Qy   457 PSLKSPPPGFDSVIARGQTEPDPAQDIELELDGQPV 492
             || :|| :|: :   ::||:
Db   531 SSL-----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 560

RESULT 13
US-08-391-259-7
; Sequence 7, Application US/08391259
; Patent No. 5621078
; GENERAL INFORMATION:
; APPLICANT: Riemen, Mark W
; APPLICANT: Stirdivant, Steven M
; TITLE OF INVENTION: Modified PE40
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: U.S.
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,259
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/120,698
; FILING DATE:
; APPLICATION NUMBER: US/07/879,037
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Grassler, Frank P
; REGISTRATION NUMBER: 31,164
; REFERENCE/DOCKET NUMBER: 17879IA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3462
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-391-259-7

Query Match   3.6%; Score 100.5; DB 1; Length 420;
Best Local Similarity   21.2%; Pred. No. 0.21;
Matches   67; Conservative   38; Mismatches 126; Indels   85; Gaps 14;

Qy   215 MPLGKLTKQQIARGFEALE----------------------------ALEEAMKNPTGDGQSL 249
             :|| ||: | ||:| | |
Db   74 LPLETLTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWNQVDQVIRNALASP-GSGGDL 132

Qy   250 EELSSCFYTVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEEKVE 309
             : | : | : || ::| | ::| :| ||
Db   133 GEAI----------------REQPE-------QAR--LALTLAAAESERFVRQGTGNDEAGA- 169

Qy   310 EVPHPLDRDYQLLRCQLQL-----LDSG----ESEYKAIQTYLKQTGNSYRCPNLRHVW 359
             : | : | |: :| | :||:||
Db   170 ----ANADVVSLTCPVAAGECAGPADSGDALLERNYPTEAEFLGDGGDVSFSTRGTQNW 224

Qy   360 KVNREGEGDRFQAHSKLGNRRLL--WHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFA 416
             : | || :||:|| :||:| |:|| :: |:| :| :| || :||
Db   225 TVER------LLQAHRQLEERGYVFVGYHGTFLEAAQSIVFGGVRARSQDLDAIWRGFYIA 279

Qy   417 SENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTE 476

```
Db  280 GDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPRSSL----PGFYRT-SLTLAA 327
         :: ||       : ::  ||       |||
Qy  477 PDPAQDIELELDGQPV 492
        |:|:||  ||||
Db  328 PEAAGEVE-RLIGHPL 342
```

RESULT 14
US-08-839-425-7
; Sequence 7, Application US/08839425
; Patent NO. 5912322
; GENERAL INFORMATION:
;   APPLICANT: Riemen, Mark W
;   APPLICANT: Stirdivant, Steven M
;   TITLE OF INVENTION: Modified PE40
;   NUMBER OF SEQUENCES: 11
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Merck & Co., Inc.
;   STREET: 126 Lincoln Avenue
;   CITY: Rahway
;   STATE: New Jersey
;   COUNTRY: U.S.
;   ZIP: 07065
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: Power MacIntosh 6.0.1
;   SOFTWARE: Microsoftword 6.0.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/839,425
;   FILING DATE:
;   CLASSIFICATION: 530
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Grassler, Frank P
;   REGISTRATION NUMBER: 31,164
;   REFERENCE/DOCKET NUMBER: 17879IA
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (908) 594-3462
;   TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 420 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-839-425-7

Query Match   3.6%; Score 100.5; DB 2; Length 420;
Best Local Similarity 21.2%; Pred. No. 0.21;
Matches 67; Conservative 38; Mismatches 126; Indels 85; Gaps 14;

```
Qy  215 MPLGKLTKQQIARGFEALE-----------------ALEEAMKNPTGDGQSL 249
         |||      ||:|| ||              |: ::|  ||
Db   74 LPLETLTRHRQPRGWEQLEQAGIYPQRLVALYLAARLSWNQVDQVIRNALASP-GSGGDL 132
Qy  250 EELSSCFYTVIPHNFGRSRPPPINSPDVLQAKKDMLLVLADIELAQTLQAAPGEEEEKVE 309
         |:|         |   |:|   ||:
Db  133 GEAI---------------QAR--LALTLAAAESERFVRQGTGNDEAGA- 169
Qy  310 EVPHPLDRDYQLLRCQLQL----LDSG----ESEYKAIQTYLKQTGNSYRCPNLRHVW 359
         |    |||  :   |  ||          |   |            |  :|
Db  170 ----ANADVVSLTCPVAAGECAGPADSGDALLERNYPTEAAEFLGDGGDVSFSTRGTQNW 224
Qy  360 KVNREGEGDRFQAHSKLGNRLL---WHGTNVAVVAAILTSGLRIMPHSGRVGKGIYFA 416
         ||              |||::   ||       |:| |:|    :||||
Db  225 TVER----LLQAHRQLEERGYVFVGYHGTFLEAAQSIVFGGVRARSQDLDAIWRGFYIA 279
Qy  417 SENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTE 476
         |       || |   |:||:|  ||||   |       |       ||
Db  280 GDPALAYGYAQDQEPDAR------GRIRNGALLRVYVPRSSL----PGFYRT-SLTLAA 327
```

```
Db  477 PDPAQDIELELDGQPV 492
        |:|:||  ||||
Db  328 PEAAGEVE-RLIGHPL 342
```

RESULT 15
US-09-336-447A-15
; Sequence 15, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
;   APPLICANT: HANSEN, ERIC J.
;   APPLICANT: AEBI, CHRISTOPH
;   APPLICANT: COPE, LESLIE D.
;   APPLICANT: MACIVER, ISOBEL
;   APPLICANT: FISKE, MICHAEL J.
;   APPLICANT: FREDENBURG, ROSS A.
;   TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS
;   FILE REFERENCE: AMCY:024
;   CURRENT APPLICATION NUMBER: US/09/336,447A
;   CURRENT FILING DATE: 1999-06-21
;   NUMBER OF SEQ ID NOS: 98
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 889
;   TYPE: PRT
;   ORGANISM: Moraxella catarrhalis
US-09-336-447A-15

Query Match   3.6%; Score 100.5; DB 4; Length 889;
Best Local Similarity 19.9%; Pred. No. 0.69;
Matches 101; Conservative 78; Mismatches 166; Indels 163; Gaps 25;

```
Qy    8 SVQTEGSKKQ--------RQGTEEDSFR-STAEALRAAPADNRVIRVDPSCPFSRNPGIQ 59
         ||   ::|
Db  423 SVGSEGKERQIVHVGAGRISNDSTDAVNGSQLYALAAAVDDNQY------------DIE 469
Qy   60 VHEDYDCTLNQTNIGNNNNKFYIIQLLEEGSRFFCWNRWGR-------------VGEVG 105
         ::||| |          :|
Db  470 KNQD-DIAKNQADIAKNQAD--IQTLENDVGKELLNLSGRLIDQKADIDNNIHIYELA 525
Qy  106 QSKMNHFTCLEDAKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVGQEAESQEA---- 161
         ::::::::                ::|
Db  526 QQQDQHSSDIKTLKKNVEE--------------------GLLELSGHLIDQKADLTK 562
Qy  162 VVKALSPQVDSGPVRTVVKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLT 221
         :|         :|                     :|
Db  563 DIKALESNVEEG-----------LLDLSGRLIDQ--KADIAQNQANIQDL---AAYNELQ 606
Qy  222 KQQIARGFEALEALEEAMKNPTGDGQSLEELSSCFYTVIPHNFGRSRPPPI--------- 272
         |:|   |     |||        |
Db  607 DQYAQKQTEAIDALNKASSENT--QNIEDLAA--YNELQDAYAKQQTEAIDALNKASSE 661
Qy  273 NSPDVLQAKKDMLLVLADI-ELAQ--------------TLQAAPGEEEKVEVPHPLDRDYQ- 320
         |:  :|||  :| ||:     |||            :: ::
Db  662 NTQNIAKNQADIANNINNIYELAQQQDQHSSDIKTLAKASAANTDRIAKNKADADASFET 721
Qy  321 LLRCQLQLLD---------------SGESEYKAIQTYLKQTGNSYRCPNLRHVW 359
         |:|      |             |||::|
Db  722 LTKNQNTLIEKDKEHDKLITANKTAIDANKASADTKFAATADAITKNGNAI-TKNAKSIT 780
Qy  360 KVNREGEG--------------DRFQA-HSKLGNRLLWHGTNVAVVAAILTSGLRIM 402
         ||             |||          ||||
Db  781 DLGTKVDGFDGRVTALDTKVNAFDGRITALDSKVEN-------GMAAQAAL--SGL-FQ 829
Qy  403 PHSGGRVGKGIYFASENSKS-----AGY 425
         |||
Db  830 PYSVGKFNATAALGGYGSKSAVAIGAGY 857
```

Search completed: August 29, 2002, 07:58:21

THIS PAGE BLANK (USPTO)

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 29, 2002, 07:52:23 ; Search time 124.84 Seconds
                                        (without alignments)
                                        507.146 Million cell updates/sec

Title:           US-09-701-586B-2
Perfect score:   2998
Sequence:        1 MAARRRRSTGGGRARALNES.........PNQVRMRYLLKVQFNFLQLW 570

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters:          747574

Minimum DB seq length:  0
Maximum DB seq length:  2000000000
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :       A_Geneseq_032802:*
     1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
     2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
     3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
     4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
     5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
     6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
     7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
     8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
     9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
    10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
    11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
    12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
    13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
    14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
    15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
    16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
    17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
    18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
    19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
    20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
    21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
    22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2998	100.0	570	21	AAY51174	Human brain PARP2
2	2998	100.0	570	22	AAB11480	Human brain poly-A
3	2981.0	99.4	583	22	AAB47029	hPARP2. Homo sapi
4	2749	91.7	521	22	AAB60693	Human poly(ADP-rib
5	2732.5	91.1	534	21	AAB42909	Human ORFX ORF2673
6	2732.5	91.1	534	22	AAU29023	Human PARP-2 prote
7	2723.5	90.8	534	22	AAU29020	Human PARP-2 prote
8	2693.5	89.8	531	22	AAB93513	Human protein sequ
9	2468	82.3	522	22	AAU29022	Mouse PARP-2 prote
10	2043.5	68.2	1063	22	AAB47032	Fusion protein PAR
11	1966	65.6	379	22	AAU21687	Novel human neopla
12	1790.5	59.7	360	22	AAB47030	N-terminal fragmen
13	1518	50.6	287	22	AAB47031	C-terminal fragmen
14	1456	48.6	294	22	AAU21810	Human DNA repair a
15	1456	48.6	294	21	AAY68835	Novel human neopla
16	1180	39.4	637	21	AAY68834	The poly(ADP-ribos
17	1149.5	38.3	653	21	AAY20130	A poly(ADP-ribose)
18	1145	38.2	227	22	AAU20130	Human DNA repair a
19	1145	38.2	227	22	AAU21811	Novel human neopla
20	1012.5	33.8	1013	17	AAR99642	Poly(ADP-ribose) p
21	1010.5	33.7	1014	20	AAY33699	Human poly(ADP-rib
22	1007.5	33.6	1014	21	AAY58043	Human poly (ADP-rl
23	1007.5	33.6	1014	22	AAU29019	Human PARP-1 prote
24	1007.5	33.6	1014	22	AAB66296	Human tankyrase2 r
25	998.5	33.3	1014	21	AAY49939	Human nuclear NAD+
26	979.5	32.7	557	22	ABB66431	Drosophila melanog
27	961.5	32.1	969	21	AAY68833	A poly(ADP-ribose)
28	961.5	32.1	980	21	AAY68839	A poly(ADP-ribose)
29	911	30.4	190	22	AAU21688	Novel human neopla
30	847.5	28.3	982	20	AAY28464	Maize poly ADP-rib
31	725	24.2	533	21	AAY51175	Human brain PARP3
32	725	24.2	540	21	AAY51176	Human uterus type
33	720	24.0	533	22	AAU29021	Human PARP-3 prote
34	700.5	23.4	528	21	AAY51178	Murine PARP1 (shor
35	698	23.3	533	21	AAY51177	Murine PARP1 (long
36	637	21.2	122	21	AAG01000	Human secreted pro
37	576	19.2	1010	21	AAY68840	Fusion protein of
38	306.5	10.2	1724	21	AAY54373	cDNA sequence enco
39	306.5	10.2	1724	22	AAB51022	Human minor vault
40	304.5	10.2	1730	22	AAU33242	Novel human secret
41	299	10.0	1099	22	AAB66301	Human tankyrase2 e
42	141	4.7	991	22	AAB47023	Mouse SPANK. Mus
43	139.5	4.7	1515	22	ABG10624	Novel human diagno
44	136.5	4.6	1442	22	ABG25501	Novel human diagno
45	136.5	4.6	2408	22	ABG10631	Novel human diagno

ALIGNMENTS

```
RESULT 1
AAY51174
ID   AAY51174 standard; Protein; 570 AA.
XX
AC   AAY51174;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human brain PARP2 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI: 2000-087218/07.
DR   N-PSDB; AAZ44287.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
```

DT   02-MAR-2001 (first entry)
XX
DE   Human brain poly-ADP-ribose-polymerase protein.
XX
KW   Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
KW   cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
KW   anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
KW   antiinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
KW   2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
KW   neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
KW   Parkinson's disease; ischemic damage; microinfarction; sepsis;
KW   diabetes mellitus.
XX
OS   Homo sapiens.
XX
PN   DE19921567-A1.
XX
PD   16-NOV-2000.
XX
PF   11-MAY-1999;   99DE-1021567.
XX
PR   11-MAY-1999;   99DE-1021567.
XX
PA   (BADI) BASF AG.
XX
PI   Lubisch W, Sadowski J, Kock M, Hoeger T;
XX
DR   WPI; 2001-032983/05.
DR   N-PSDB; AAC82090.
XX
PT   Drugs for inhibiting PARP or especially homologous enzymes comprising
PT   4-substituted phthalazinone derivatives, useful e.g. for treating
PT   neurodegenerative disease, ischemic damage, tumors or diabetes -
XX
PS   Example A; Page 12-13; 14pp; German.
XX
CC   This invention describes novel 4-substituted 2H-phthalazin-1-one
CC   derivatives (I) which are used for the treatment or prophylaxis of
CC   diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;
CC   EC 2.4.2.30) activity. The products of the invention have nootropic,
CC   neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,
CC   cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,
CC   immunosuppressive, antiinflammatory, antirheumatic, antiarthritic,
CC   antidiabetic. (I) are especially used for treating or preventing
CC   neurodegenerative disease or neuronal damage (specifically associated
CC   with ischemia, trauma or massive bleeding, especially apoplexy or
CC   spinal-cranial trauma; or Alzheimer's disease, Huntington's disease or
CC   Parkinson's disease), treating or preventing ischemic damage
CC   (specifically renal damage after renal ischemia or during and after
CC   kidney transplantation or heart damage after cardiac ischemia), treating
CC   epilepsy, specifically generalized epileptic attacks (e.g. petit mal and
CC   tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lope
CC   and complex partial attacks), treating microinfarction (e.g. during and
CC   after heart valve replacement, aneurysm resectioning and heart
CC   transplantation), revascularization of critically constricted coronary
CC   arteries (e.g. after PCTA or by-pass operations) or peripheral arteries
CC   (e.g. leg arteries), treating acute myocardial ischemia and damage during
CC   or after its mechanical or drug-induced lysis and treating tumors and
CC   their metastasis, sepsis and septic shock, inflammatory and rheumatic
CC   disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)
CC   inhibit PARP (i.e. the known form designated PARP1), they especially
CC   selectively and strongly inhibit PARP homologs, specifically the homolog
CC   PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.
CC   with K_1 values of 1-20 nM) and high selectivity for PARP2 relative to
CC   PARP1 (generally by a factor of more than 5).
XX
SQ   Sequence   570 AA;

PT   conditions -
XX
PS   Claim 4; Page 52-54; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP2 protein used in the method of the invention.
XX
SQ   Sequence   570 AA;

Query Match   100.0%; Score 2998; DB 21; Length 570;
Best Local Similarity   100.0%; Pred. No. 2.6e-275;
Matches 570; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MAARRRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 maarrrrstgggraralneskrvnngntapedsspakktrrcqrqeskkmpvaggkankd 60

Qy   61 RTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDYDVMLNQTNLQFNNNKYYL 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 rtedkqdesvkalllkgkapvdpectakvgkahvycegndydvmlnqtnlqfnnnkyyl 120

Qy   121 IQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 iqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnkakeifqkkfldktknnwedrek 180

Qy   181 FEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDRVQELIKLICNVQAMEE 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 fekvpgkydmlqmdyatntqdeeetkkeeslkspllkpesqldrvqeliklicnvqamee 240

Qy   241 MMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 mmmemkyntkkaplgkltvaqikagyqslkkiedciragqhgralmeacnefytriphdf 300

Qy   301 GLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLD 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 glrtpplirtqkelsekiqllealgdieiaiklvktelqspehpldqhyrnlhcalrpld 360

Qy   361 HESYEFKVISQYLQSTHAPHSDYTMTLLDLFEVEKDGEKEAFRRDLHNRMLLWHGSRMS 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 hesyefkvisqylqsthapthsdytmtldlifevekdgekeafredlhnrmllwhgsrms 420

Qy   421 NWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 nwvgilshglriappeapigymfgkgiyfadmssksanycfasrlkntglllsevalg 480

Qy   481 QCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 qcnelleanpkaegllqgkhstkglgkmapssahfvtlngstvplgpasdtgilnpdgyt 540

Qy   541 LNYNEYIVNPNQVRMRYLLKVQFNFLQLW 570
           |||||||||||||||||||||||||||||
Db   541 lnyneyivnpnqvrmryllkvqfnflqlw 570

RESULT 2
AAB11480
ID   AAB11480 standard; Protein; 570 AA.
XX
AC   AAB11480;
XX

```
OS   Homo sapiens.
XX
PN   WO200077179-A2.
XX
PD   21-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16629.
XX
PR   16-JUN-1999; 99US-0139543.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR   WPI: 2001-025335/03.
DR   N-PSDB; AAC85303.
XX
PT   New human poly(ADP-ribose) polymerase for treating inflammatory,
PT   neurological, cardiovascular, or neoplastic tissue growth disorders,
PT   such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
PT   metastasis -
XX
PS   Claim.3; Page 94-95; 129pp; English.
XX
CC   This sequence represents human poly(ADP-ribose) polymerase (hPARP2).
CC   This protein causes the covalent addition of polymers of ADP-ribose
CC   to protein targets.  hPARP2 activity is induced in many instances of
CC   oxidative stress or during inflammation where there is direct damage
CC   to the DNA.  hPARP2 may be used to identify antagonists which
CC   may be used to treat a human having a disorder mediated by PARP2
CC   activity, such as, inflammatory, neurological, cardiovascular,
CC   or neoplastic tissue growth disorders, e.g. ischemic stroke,
CC   hemorrhagic shock, myocardial ischemia or infarction,
CC   transplantation, cerebral vasospasm; rheumatoid, osteo- or gouty
CC   arthritis, spondylitis; Behcet's disease; sepsis, septic or endotoxic
CC   shock, gram negative or positive sepsis, toxic shock syndrome; multiple
CC   organ injury syndrome secondary to septicemia, trauma, or hemorrhage;
CC   allergic or vernal conjunctivitis, uveitis, thyroid-associated
CC   ophthalmopathy; eosinophilic granuloma; asthma, chronic bronchitis,
CC   allergic rhinitis, ARDS, chronic obstructive pulmonary disease,
CC   silicosis, pulmonary sarcoidosis, pleurisy, alveolitis, vasculitis,
CC   pneumonia, bronchiectasis, pulmonary oxygen toxicity; reperfusion
CC   injury of the myocardium, brain or extremities; cystic fibrosis; keloid
CC   formation, scar tissue formation; atherosclerosis; systemic lupus
CC   erythematosus, autoimmune thyroiditis, multiple sclerosis; Reynaud's
CC   syndrome; graft versus host disease, allograft rejection; chronic
CC   glomerulonephritis; inflammatory bowel disease, Crohn's disease,
CC   ulcerative colitis, necrotizing enterocolitis; inflammatory dermatoses,
CC   contact or atopic dermatitis, psoriasis, urticaria, fever and myalgias
CC   due to infection: meningitis, encephalitis, and brain and spinal cord
CC   injury due to minor trauma; Sjogren's syndrome; diseases involving
CC   leukocyte diapedesis; alcoholic hepatitis; bacterial pneumonia;
CC   antigen-antibody complex mediated diseases; hypovolemic shock; Type 1
CC   diabetes mellitus; acute and delayed hypersensitivity; disease states
CC   due to leukocyte dyscrasia and metastasis; thermal injury; granulocyte
CC   transfusion associated syndromes; and cytokine-induced toxicity.
CC   hPARP2 and antibodies to it, can also be used to diagnose these
CC   conditions.
XX
SQ   Sequence   583 AA;

Query Match             99.4%;  Score 2981.5;  DB 22;  Length 583;
Best Local Similarity   97.8%;  Pred. No. 1e-273;
Matches  570;  Conservative   0;  Mismatches   0;  Indels  13;  Gaps  1;

Qy          1 MAARRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 maarrrstgggraralneskrvnngntapedsspakktrrcqrqeskkmpvaggkankd 60

Qy         61 RTEDKQD------------ESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLN 107
              ||||||||                    ||||||||||||||||||||||||||||||
Db         61 rtedkqdmpgrswaskrvsesvkalllkgkapvdpectakvgkahvycegndvydvmln 120
```

```
Qy          1 MAARRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 maarrrstgggraralneskrvnngntapedsspakktrrcqrqeskkmpvaggkankd 60

Qy         61 RTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQFNNNKYYL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 rtedkqdesvkalllkgkapvdpectakvgkahvycegndvydvmlnqfnnnkyyl 120

Qy        121 IQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKFLDKTKNNWEDREK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 iqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnkakeifqkfldktknnwedrek 180

Qy        181 FEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 fekvpgkydmlqmdyatntqdeeetkkeeslksplkpesqldlrvqeliklicnvqamee 240

Qy        241 MMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 mmmemkyntkkaplgkltvaqikagyqslkkiedciragqhgralmeacnefytriphdf 300

Qy        301 GLRTPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 glrtpplirtqkelsekiqllealgdieiaiklvktelqspehpldqhyrnlhcalrpld 360

Qy        361 HESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLHNRMLLWHGSRMS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 hesyefkvisgylqsthapthsdytmtlldlfevekdgekeafredlhnrmllwhgsrms 420

Qy        421 NWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSSANYCFASRLKNTGLLLLSEVALG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 nwvgilshglriappeapitgymfgkgiyfadmsskssanycfasrlkntglllllsevalg 480

Qy        481 QCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 qcnelleanpkaegllqgkhstkglgkmapssahfvtlngstvplgpasdtgilnpdgyt 540

Qy        541 LNYNEYIVYNPNQVRMRYLLKVQFNFIQLW 570
              ||||||||||||||||||||||||||||||
Db        541 lnyneyivynpnqvrmryllkvqfnfiqlw 570

RESULT  3
AAB47029
ID   AAB47029 standard; Protein; 583 AA.
XX
AC   AAB47029;
XX
DT   29-MAR-2001  (first entry)
XX
DE   hPARP2.
XX
KW   Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
KW   inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
KW   infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
KW   gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
KW   endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
KW   toxic shock syndrome; multiple organ injury syndrome; vasculitis;
KW   hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
KW   eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
KW   chronic obstructive pulmonary disease; silicosis; reperfusion injury;
KW   pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
KW   bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
KW   scar tissue formation; atherosclerosis; systemic lupus erythematosus;
KW   autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
KW   graft versus host disease; allograft rejection; cystic fibrosis;
KW   chronic glomerulonephritis; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
KW   inflammatory dermatosis; contact dermatitis; atopic dermatitis;
KW   psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
KW   Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
KW   hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
KW   leukocyte dyscrasia; thermal injury; cytokine-induced toxicity.
XX
```

```
CC   This sequence represents a novel human poly(ADP-ribose)
CC   synthetase, sbhPARS2. The invention also relates to fragments,
CC   variants and sequences with at least 95% identity to the sbhPARS2
CC   protein or nucleotide sequence; expression systems and host cells
CC   comprising an sbhPARS2 nucleic acid sequence; the recombinant expression
CC   of sbhPARS2; and an antibody specific for sbhPARS2. sbhPARS2 proteins
CC   and nucleotides are useful as vaccines for inducing an immunological
CC   response in a mammal. The sbhPARS2 protein is useful for identifying
CC   compounds which inhibit or stimulate its activity or expression level.
CC   Such agonists and antagonists of sbhPARS2 are useful for treating human
CC   diseases including ischaemia and ischaemic tissue injury (e.g., cerebral
CC   and cardiac ischaemia, myocardial infarction, stroke), inflammation,
CC   autoimmune disease (e.g. diabetes, multiple sclerosis) and
CC   neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's
CC   disease). sbhPARS2 nucleic acids are useful as diagnostic reagents for
CC   detecting mutations in the associated gene; as hybridisation probes
CC   to isolate full-length sbhPARS2 cDNAs and sbhPARS2 genomic clones; and
CC   for chromosome localisation studies. The sbhPARS2 protein is also useful
CC   as an immunogen to produce antibodies for therapeutic use. sbhPARS2
CC   proteins, nucleotides and antibodies are also useful in screening methods
CC   for detecting the effect of added compounds on the production of mRNA and
CC   protein in cells.
XX
SQ   Sequence   521 AA;
```

```
Query Match             91.7%;  Score 2749;  DB 22;  Length 521;
Best Local Similarity   100.0%;  Pred. No. 9.9e-252;
Matches  521;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   50 MPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQT 109
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 mpvaggkankdrtedkqdesvkalllkgkapvdpectakvgkahvycegndvydvmlnqt 60

Qy  110 NLQFNNKKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLD 169
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 nlqfnnkkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnkakeifqkkfld 120

Qy  170 KTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQELI 229
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslkspLkpesqldlrvqeli 180

Qy  230 KLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEAC 289
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 klicnvqameemmemkyntkkaplgkltvaqikagyqslkkiedciragqhgralmeac 240

Qy  290 NEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHY 349
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 nefytriphdfglrtpplirtqkelsekiqllealgdieiaiklvktelqspehpldqhy 300

Qy  350 RNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLHN 409
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 rnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlfevekdgekeafredlhn 360

Qy  410 RMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNT 469
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 rmllwhgsrmsnwvgilshglriappeapitgymfgkgiyfadmsssksanycfasrlknt 420

Qy  470 GLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPAS 529
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 glllllsevalgqcnelleanpkaegllqgkhstkglgkmapssahfvtlngstvplgpas 480

Qy  530 DTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNFLQLW 570
         |||||||||||||||||||||||||||||||||||||||||
Db  481 dtgilnpdgytlnyneyivynpnqvrmryllkvqfnflqlw 521

RESULT 5
AAB42909
ID   AAB42909 standard; Protein; 534 AA.
XX
AC   AAB42909;
XX
```

```
Qy  108 QTNLQFNNKKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKF 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 qtnlqfnnkkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnkakeifqkkf 180

Qy  168 LDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQE 227
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslkspLkpesqldlrvqe 240

Qy  228 LKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALME 287
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 lklicnvqameemmemkyntkkaplgkltvaqikagyqslkkiedciragqhgralme 300

Qy  288 ACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQ 347
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 acnefytriphdfglrtpplirtqkelsekiqllealgdieiaiklvktelqspehpldq 360

Qy  348 HYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDL 407
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 hyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlfevekdgekeafredl 420

Qy  408 HNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLK 467
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 hnrmllwhgsrmsnwvgilshglriappeapitgymfgkgiyfadmsssksanycfasrlk 480

Qy  468 NTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGP 527
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ntglllsevalgqcnelleanpkaegllqgkhstkglgkmapssahfvtlngstvplgp 540

Qy  528 ASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNFLQLW 570
         |||||||||||||||||||||||||||||||||||||||||||
Db  541 asdtgilnpdgytlnyneyivynpnqvrmryllkvqfnflqlw 583

RESULT 4
AAB60693
ID   AAB60693 standard; Protein; 521 AA.
XX
AC   AAB60693;
XX
DT   11-MAY-2001 (first entry)
XX
DE   Human poly(ADP-ribose) synthetase sbhPARS2.
XX
KW   Human; poly(ADP-ribose) synthetase; sbhPARS2; vaccine; drug screening;
KW   ischaemic disorder; cerebral ischaemia; cardiac ischaemia;
KW   myocardial infarction; stroke; inflammation; autoimmune disease;
KW   diabetes; multiple sclerosis; neurodegenerative disease;
KW   Parkinson's disease; Alzheimer's disease; chromosome localisation.
OS   Homo sapiens.
XX
XX
PN   WO200112645-A1.
XX
PD   22-FEB-2001.
XX
PF   10-AUG-2000; 2000WO-US21775.
XX
PR   12-AUG-1999; 99US-0373441.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Barone F, Field J, Kabnick KS, Li X, McQueney MS, Zhu Y;
XX
DR   WPI; 2001-211196/21.
DR   N-PSDB; AAF59996.
XX
PT   New human polypeptide of the polyADPribose synthetase family for use in
PT   diagnostic assays and for screening modulators used for preventing and
PT   treating inflammation, autoimmune disease and Alzheimers disease -
XX
PS   Claim 1; Page 21; 30pp; English.
XX
```

```
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2673 polypeptide sequence SEQ ID NO:5346.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;  99US-0127607.
PR   02-APR-1999;  99US-0127636.
PR   05-APR-1999;  99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA, Leach M;
XX
DR   WPI; 2000-602362/57.
DR   N-PSDB; AAC77118.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 11; Page 4522-4524; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence, of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence   534 AA;

Query Match             91.1%;  Score 2732.5;  DB 21;  Length 534;
Best Local Similarity   97.6%;  Pred. No. 3.8e-250;
Matches  521;  Conservative   0;  Mismatches   0;  Indels  13;  Gaps   1;
```

```
Qy         50 MPVAGGKANKDRTEDKQD----------ESVKALLLKGKAPVDPECTAKVGKAHVYC 96
              ||||||||||||||||||              |||||||||||||||||||||||
Db          1 mpvaggkankdrtedkqdgmpgrswaskrvsesvkalllkgkapvdpectakvgkahvyc 60

Qy         97 EGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNL 156
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 egndydvdvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnl 120

Qy        157 NKAKEIFQKFKLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLK 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 nkakeifqkfkldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplk 180

Qy        217 PESQLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCI 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 pesqldlrvqeliklicnvqameemmmemkyntkkaplgkltvaqikagyqslkkiedci 240

Qy        277 RAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKT 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ragqhgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieiaiklvkt 300

Qy        337 ELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYIQSTHAPTHSDYTMTLLDLFEVEK 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 elqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlfevek 360

Qy        397 DGEKAFREDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGVMFGKGIYFADMSSK 456
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 dgekeafredlhnrmllwhgsrmsnwvgilshglrlappeapitgymfgkglyfadmnssk 420

Qy        457 SANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFV 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 sanycfasrlkntglllsevalggcnelleanpkaeglqgkhstkglgkmapssahfv 480

Qy        517 TLNGSTVPLGPASDTGILNPDGYTLNVNEYIVYNPNQVRMRYLLKVQFNFLQLW 570
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 tlngstvplgpasdtgilnpdgytlnyneyivynpnqvrmryllkvqfnflqlw 534
```

```
RESULT    6
AAU29023
ID   AAU29023 standard; Protein; 534 AA.
XX
AC   AAU29023;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human PARP-2 protein #2.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45684.
XX
PT   Antisense compound useful for treating hyperproliferative,
```

PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 16; Page 125-127; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cdNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence   534 AA;

Query Match   91.1%;   Score 2732.5;   DB 22;   Length 534;
Best Local Similarity 97.6%;   Pred. NO. 3.8e-250;
Matches 521; Conservative   0; Mismatches   0; Indels 13; Gaps 1;

Qy   50 MPVAGGKANKDRTEDKQD---------ESVKALLLKGKAPVDPECTAKVGKAHVYC 96
          ||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1 mpvaggkankdrtedkqdgmpgrswaskrvsesvkalllkgkapvdpectakvgkahvyc 60

Qy   97 EGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNL 156
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 egndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnl 120

Qy 157 NKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLK 216
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 nkakeifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplk 180

Qy 217 PESQLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCI 276
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 pesqldlrvqeliklicnvqameemmmemkyntkkaplgkltvaqikagyqslkkiedci 240

Qy 277 RAQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKT 336
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 raqhgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieiaiklvkt 300

Qy 337 ELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEK 396
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 elqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlfevek 360

Qy 397 DGEKEAFREDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSK 456
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 dgekeafredlhnrmllwhgsrmsnwvgilshglriappeapitgymfgkgiyfadmssk 420

Qy 457 SANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGRMAPSSAHFV 516
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 sanycfasrlkntglllsevalgqcnelleanpkaeglqgkhstkglgkmapssahfv 480

Qy 517 TLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNFLQLW 570
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 tlngstvplgpasdtgilnpdgytlnyneyivynpnqvrmryllkvqfnflqlw 534

RESULT 7
AAU29020
ID   AAU29020 standard; Protein; 534 AA.
XX
AC   AAU29020;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human PARP-2 protein #1.

XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45586.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 102-104; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cdNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence   534 AA;

Query Match   90.8%;   Score 2723.5;   DB 22;   Length 534;
Best Local Similarity 97.4%;   Pred. No. 2.7e-249;
Matches 520; Conservative   0; Mismatches   0; Indels   1; Gaps 1;

Qy   50 MPVAGGKANKDRTEDKQD---------ESVKALLLKGKAPVDPECTAKVGKAHVYC 96
          ||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1 mpvaggkankdrtedkqdgmpgrswaskrvsesvkalllkgkapvdpectakvgkahvyc 60

Qy   97 EGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNL 156
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 egndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnl 120

Qy 157 NKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLK 216
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 nkakeifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplk 180

Qy 217 PESQLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCI 276
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 pesqldlrvqeliklicnvqameemmmemkyntkkaplgkltvaqikagyqslkkiedci 240

Qy 277 RAQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKT 336
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 raqhgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieiaiklvkt 300

```
Qy  337  ELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEK  396
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  elqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlfevek  360

Qy  397  DGEKEAFREDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSK  456
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  dgekeafredlhnrmllwhgsrmsnwvgilshglriahpeapitgymfgkgiyfadmssk  420

Qy  457  SANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFV  516
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  sanycfasrlkntglllllsevalgqcnelleanpkaegllqgkhstkglgkmapssahfv  480

Qy  517  TLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNFLQLW  570
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  tlngstvplgpasdtgilnpdgytlnyneyivynpnqvrmryllkvqfnflqlw  534
```

RESULT 8

```
AAB93513
ID   AAB93513 standard; Protein; 531 AA.
XX
AC   AAB93513;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:12844.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;  99JP-0248036.
PR   27-AUG-1999;  99JP-0300253.
PR   11-JAN-2000;  2000JP-0118776.
PR   02-MAY-2000;  2000JP-0183767.
PR   09-JUN-2000;  2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 12844; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   531 AA;
```

```
Query Match          89.8%;    Score 2693.5; DB 22;   Length 531;
Best Local Similarity 97.4%;    Pred. No. 1.9e-246;
Matches  517;  Conservative  1;  Mismatches  1;  Indels  13;  Gaps  1;

Qy    1  MAARRRSTGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD   60
         |||||||||||||||||||||||||||             ||||||||||||||||||||
Db    1  maarrrstggraralneskrvnngntapedsspakktrrcqrgqeskkmpvaggkankd   60

Qy   61  RTEDKQD----------ESVKALLLKGAKPVDPECTAKVGKAHVYCEGNDVYDVMLN  107
         |||||||            ||||||||||||||||||||||||||||||||||||||||
Db   61  rtedkqdgmpgrswaskrvsesvkalllkgakpvdpectakvgkahvycegndvydvmln  120

Qy  108  QTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKF  167
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  qtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnkakeifqkkf  180

Qy  168  LDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQE  227
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  ldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpesqldlrvqe  240

Qy  228  LIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALME  287
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  liklicnvqameemmemkyntkkaplgkltvaqikagyqslkkiedciragqhgralme  300

Qy  288  ACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQ  347
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  acnefytriphdfglrtpplirtqkelsekiqllealgdieiaiklvktelqspehpldq  360

Qy  348  HYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDL  407
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  hyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlfevekdgekeafredl  420

Qy  408  HNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLK  467
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  hnrmllwhgsrmsnwvgilshglriappeapitgymfgkgiyfadmssksanycfasrlk  480

Qy  468  NTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTL  518
         :|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  htglllllsevalgqcnelleanpkaegllqgkhstkglgkmapssahfvtl  531
```

RESULT 9

```
AAU29022
ID   AAU29022 standard; Protein; 522 AA.
XX
AC   AAU29022;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Mouse PARP-2 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Mus musculus.
XX
PN   WO20016495S-A1.
XX
PD   07-SEP-2001.
```

us-09-701-586b-2.rag

```
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45597.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 109-111; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase) plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence  522 AA;
```

```
Query Match             82.3%;  Score 2468; DB 22;  Length 522;
Best Local Similarity   87.6%;  Pred. No. 4.7e-225;
Matches  465; Conservative 35; Mismatches  21; Indels  10; Gaps  4;

Qy    40 RRCQROESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGN 99
            ||||||| ||||:|:||||| ||||| :||||: |  |:||||||||| |||||||||||
Db     2 rtcqr---kgpmaggk-dadrtkdnrd-svktlllkgkapvdpecaakigkahvycegd 55

Qy   100 DVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKA 159
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||:|||
Db    56 dvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkvgvtcsgdlnka 115

Qy   160 KEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPES 219
           ||||||||||||||||||||||||| |||:|:|| |||:|||:|| |:||: |||||
Db   116 keifqkfldktknnwedrenfekvpgkydmlqmdyaastqdesktkeeet---1kpes 171

Qy   220 QLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAG 279
           |||||||||||||||||||||| ||||||||||||||||||:||||||||||||||
Db   172 qldlrvqeliklicnvqtmeemniemkydtkrapigkltvaqikagyqslkkiedcirag 231

Qy   280 QHGRALMEACNEFYTRIPHDFGLRTPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQ 339
           |||:|:||||||||||||||||||||| ||:||||:||:::||||:||||||||||:| |
Db   232 qhgralveacnefytriphdfgls ippvirtekelsdkvkllealgdieialklvkserq 291

Qy   340 SPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE 399
           ||||||||||||||||||| |||:|||:| ||:||:||||||:|||:||||:||||:|||:||
Db   292 glehpldqhyrnlhcalrpldhesnefkvisqylqsthapthkdytmtlldvfevekege 351

Qy   400 KEAFREDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPTTGYMFGKGIYFADMSSKSAN 459
           ||||||||||||||||||||||||||||||||| :||||||:||||||||||:|||||:||
Db   352 keafredlpnrmllwhgsrisnwvgilshglrvappeapitgymfgkgiyfadmssksan 411

Qy   460 YCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLN 519
           ||||||||||||||| ||||||||||||||||||:||||:|:||:|||:||::|||||||
Db   412 ycfasrlkntglllsevalgqcnelleanpkaqglllrgkhstkgmgkmapspahfitln 471

Qy   520 GSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNFLQLW 570
```

```
Db   472 gstvplgpasdtgilnpegytlnynefivyspnqvmryllkiqfnflqlw 522
```

```
RESULT 10
AAB47032
ID   AAB47032 standard; Protein; 1063 AA.
XX
AC   AAB47032;
XX
DT   29-MAR-2001 (first entry)
XX
DE   Fusion protein PARP1A/PARP2B.
XX
KW   Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
KW   inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
KW   infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
KW   gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
KW   endotoxic shock; gram negative sepsis; multiple organ injury syndrome; vasculitis;
KW   toxic shock syndrome; gram positive sepsis; trauma;
KW   hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
KW   eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
KW   chronic obstructive pulmonary disease; silicosis; reperfusion injury;
KW   pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
KW   bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
KW   scar tissue formation; atherosclerosis; systemic lupus erythematosus;
KW   autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
KW   graft versus host disease; allograft rejection; cystic fibrosis;
KW   chronic glomerulonephritis; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
KW   inflammatory dermatosis; contact dermatitis; atopic dermatitis;
KW   psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
KW   Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
KW   hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
KW   leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;
KW   expressed sequence tag; EST; RACE; PCR; amplify; primer;
KW   polymerase chain reaction.
XX
OS   Synthetic.
XX
PN   WO20007179-A2.
XX
PD   21-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16629.
XX
PR   16-JUN-1999; 99US-0139543.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  McElligott DL;
XX
DR   WPI; 2001-025335/03.
DR   N-PSDB; AAC85341.
XX
PT   New human poly(ADP-ribose) polymerase for treating inflammatory,
PT   neurological, cardiovascular, or neoplastic tissue growth disorders,
PT   such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
PT   metastasis -
XX
PS   Example 3; Page 117-120; 129pp; English.
XX
CC   This sequence represents the fusion protein PARP1A/PARP2B. This
CC   protein contains amino acids 1-662 of hPARP1 fused upstream of
CC   amino acids 230-583 of hPARP2. The cDNA encoding this sequence was
CC   amplified using the primer sequences given in AAC85321-40 and
CC   AAC85342-51. The fusion protein coding sequence was cloned in a
CC   baculovirus expression vector for the expression of the fusion protein.
CC   The protein of the invention, hPARP2, causes the covalent addition of
CC   polymers of ADP-ribose to protein targets. hPARP2 activity is induced
CC   in many instances of oxidative stress or during inflammation where
CC   there is direct damage to the DNA. hPARP2 may be used to identify
CC   antagonists which may be used to treat a human having a disorder
```

```
CC   mediated by PARP2 activity, such as, inflammatory, neurological,
CC   cardiovascular, or neoplastic tissue growth disorders. hPARP2 and
CC   antibodies to it, can also be used to diagnose these conditions.
XX
SQ   Sequence 1063 AA;

Query Match             68.2%;  Score 2043.5;  DB 22;  Length 1063;
Best Local Similarity   75.7%;  Pred. No. 3.1e-184;
Matches    411;  Conservative   32;  Mismatches   81;  Indels   19;  Gaps    6;

Qy         29 APEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLKGKAPVDPECTAK 88
              ||  : :  :  |   |  |    :: :   ||   |   |  | |   ||| | |
Db        530 aprgksgaalskksksqqv--------keeginksekirmk-ltlkggaavdpd-sgl 575

Qy         89 VGKAHVYCEGNDVYDVMLNQTNLQFPNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHS 148
              |  : :|     :  | |: : : |:|||||| |||||||| : : :||||||||| :
Db        576 ehsahvlekggkvfsatlglvdivkgtnsyyklqlleddkenrywifrswgrvgtvigsn 635

Qy        149 LVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEETKK- 207
              :        | :  :| |   |: :: : :    :|   :::: :|   : :|::|
Db        636 kleqmpskedaiehfmklyeektgnawhsk-nftkypkkfypleidyg--qdeeavkkl 691

Qy        208 EESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQ 267
              :     ||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        692 tvnpgtkrypesqldlrvgeliklicnvqameemmemkyntkkaplgkltvaqikagyq 751

Qy        268 SLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTQKELSEKIQLLEALGDI 327
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 slkkiedclragqhgralmeacnefytriphdfglrtqkelsekiqllealgdi 811

Qy        328 EIAILKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMT 387
              |:||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 elailvktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmt 871

Qy        388 LLDLFEVEKDGEKEAFREDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKG 447
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        872 lldlfevekdgekeafredlhnrmllwhgsrmsnwvgilshglriappeapitgymfgkg 931

Qy        448 IYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGK 507
              ||||||| ||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        932 iyfadmsssksanycfasrlkntgllllsevalggcnelleanpkaegllqgkhstkglgk 991

Qy        508 MAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNFL 567
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        992 mapssahfvtlngstvplgpasdtgilnpdgytlnyneyivynpnqvrmryllkvqfnfl 1051

Qy        568 QLW 570
              |||
Db       1052 qlw 1054

RESULT 11
AAU21687
ID   AAU21687 standard; Protein; 379 AA.
XX
AC   AAU21687;
XX
DT   06-DEC-2001  (first entry)
XX
DE   Novel human neoplastic disease associated polypeptide #120.
XX
KW   Human; neoplastic disease associated polypeptide; cancer;
KW   hyperproliferative disorder; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   neuroprotective; cytostatic; anti inflammatory; vasotropic.
XX
OS   Homo sapiens.
XX
PN   WO200155163-A1.
XX
PD   02-AUG-2001.

XX
PF   17-JAN-2001; 2001WO-US01358.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
```

```
PR    29-SEP-2000;    2000US-0236369.
PR    29-SEP-2000;    2000US-0236370.
PR    02-OCT-2000;    2000US-0236802.
PR    02-OCT-2000;    2000US-0237037.
PR    02-OCT-2000;    2000US-0237038.
PR    02-OCT-2000;    2000US-0237039.
PR    02-OCT-2000;    2000US-0237040.
PR    13-OCT-2000;    2000US-0239935.
PR    20-OCT-2000;    2000US-0239937.
PR    20-OCT-2000;    2000US-0240960.
PR    20-OCT-2000;    2000US-0241221.
PR    20-OCT-2000;    2000US-0241785.
PR    20-OCT-2000;    2000US-0241786.
PR    20-OCT-2000;    2000US-0241787.
PR    20-OCT-2000;    2000US-0241808.
PR    20-OCT-2000;    2000US-0241809.
PR    20-OCT-2000;    2000US-0241826.
PR    01-NOV-2000;    2000US-0244617.
PR    08-NOV-2000;    2000US-0246474.
PR    08-NOV-2000;    2000US-0246475.
PR    08-NOV-2000;    2000US-0246477.
PR    08-NOV-2000;    2000US-0246478.
PR    08-NOV-2000;    2000US-0246523.
PR    08-NOV-2000;    2000US-0246524.
PR    08-NOV-2000;    2000US-0246525.
PR    08-NOV-2000;    2000US-0246526.
PR    08-NOV-2000;    2000US-0246527.
PR    08-NOV-2000;    2000US-0246528.
PR    08-NOV-2000;    2000US-0246532.
PR    08-NOV-2000;    2000US-0246609.
PR    08-NOV-2000;    2000US-0246610.
PR    08-NOV-2000;    2000US-0246611.
PR    08-NOV-2000;    2000US-0246613.
PR    17-NOV-2000;    2000US-0249207.
PR    17-NOV-2000;    2000US-0249209.
PR    17-NOV-2000;    2000US-0249210.
PR    17-NOV-2000;    2000US-0249211.
PR    17-NOV-2000;    2000US-0249212.
PR    17-NOV-2000;    2000US-0249213.
PR    17-NOV-2000;    2000US-0249214.
PR    17-NOV-2000;    2000US-0249215.
PR    17-NOV-2000;    2000US-0249216.
PR    17-NOV-2000;    2000US-0249217.
PR    17-NOV-2000;    2000US-0249218.
PR    17-NOV-2000;    2000US-0249244.
PR    17-NOV-2000;    2000US-0249245.
PR    17-NOV-2000;    2000US-0249264.
PR    17-NOV-2000;    2000US-0249265.
PR    17-NOV-2000;    2000US-0249297.
PR    17-NOV-2000;    2000US-0249299.
PR    17-NOV-2000;    2000US-0249300.
PR    01-DEC-2000;    2000US-0250160.
PR    01-DEC-2000;    2000US-0250391.
PR    08-DEC-2000;    2000US-0251030.
PR    05-DEC-2000;    2000US-0251988.
PR    05-DEC-2000;    2000US-0256719.
PR    06-DEC-2000;    2000US-0251479.
PR    08-DEC-2000;    2000US-0251856.
PR    08-DEC-2000;    2000US-0251868.
PR    08-DEC-2000;    2000US-0251869.
PR    08-DEC-2000;    2000US-0251989.
PR    08-DEC-2000;    2000US-0251990.
PR    11-DEC-2000;    2000US-0254097.
PR    05-JAN-2001;    2001US-0259678.
XX
PA    (HUMA-) HUMAN GENOME SCI INC.
XX
PI    Rosen CA,  Barash SC,  Ruben SM;
XX
DR    WPI; 2001-465558/50.
DR    N-PSDB; AAS34886.
```

```
XX
PT    Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT    diagnose diseases or disorders associated with aberrant expression or
PT    activity of polypeptides, and for treating cancers, rheumatoid
PT    arthritis -
XX
PS    Claim 11; SEQ ID No 414; 687pp; English.
XX
CC    The present invention relates to the isolation of novel human neoplastic
CC    disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
CC    sequences encoding for these polypeptides. The sequences of the
CC    invention are useful in the diagnosis, treatment, prevention and/or
CC    prognosis of disorders involving neoplastic disease such as
CC    hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC    cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC    astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC    also be useful for treating other disorders such as neural disorders,
CC    immune system disorders, muscular disorders, reproductive disorders,
CC    gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC    and renal disorders. The polynucleotide sequences of the invention are
CC    also useful in gene therapy. AAU21568-AAU21851 represent the novel human
CC    neoplastic disease associated polypeptides of the invention.
CC    Note: The sequence data for this patent did not form part of the printed
CC    specification, but was obtained in electronic format directly from WIPO
CC    at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ    Sequence    379 AA;

Query Match             65.6%;  Score 1966;  DB 22;  Length 379;
Best Local Similarity   98.9%;  Pred. No. 1.3e-177;
Matches 375;  Conservative   1;  Mismatches  3;  Indels   0;  Gaps    0;

Qy        192 QMDYATNTQDEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMEMKYNTKK 251
                : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 exdyatntqdeeetkkeeslksplkpesqldlrvqeliklicnvqameemmemkyntkk 60

Qy        252 APLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQ 311
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aplgkltvaqikagyqslkkiedciragqhgxalmeacnefytriphdfgirtpplirtq 120

Qy        312 KELSEKIQLLEALGDIETAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQ 371
                ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 kelxekiqllealgdieiaiklvktelqspehpldqhyrnlhcalrpldhesyefkvisq 180

Qy        372 YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLHNRMLLWHGSRMSNWVGILSHGLR 431
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ylqsthapthsdytmtlldlfevekdgekeafredlhnrmllwhgsrmsnwvgilshglr 240

Qy        432 IAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPK 491
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 iappeapitgymfgkgiyfadmssksanycfasrlkntglllseevalgqcnelleanpk 300

Qy        492 AEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNP 551
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 aegllqgkhstkglgkmapssahfvtlngstvplgpasdtgilnpdgytlnyneyivynp 360

Qy        552 NQVRMRYLLKVQFNFLQLW 570
                |||||||||||||||||||
Db        361 nqvrmryllkvqfnflqlw 379

RESULT 12
AAB47030
ID    AAB47030 standard;  Protein; 360 AA.
XX
AC    AAB47030;
XX
DT    29-MAR-2001  (first entry)
XX
DE    N-terminal fragment of hPARP2.
XX
KW    Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
```

```
Db    1 maarrrrstggraralneskrvngntapedsspakktrrcqrgqeskkmpvaggkankd 60

Qy   61 RTEDKQD-------------ESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLN 107
        |||||||             ||||||||||||||||||||||||||||||||||||||||
Db   61 rtedkqdgmpgrswaskrvsesvkalllkgkapvdpectakvgkahvycegndvydvmln 120

Qy  108 QTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWRGVGKMGQHSLVACSGNLNKAKEIFQKKF 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 qtnlqfnnnkyyliqlleddaqrnfsvwmrwrgvgkmgqhslvacsgnlnkakeifqkkf 180

Qy  168 LDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQE 227
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpesqldlrvqe 240

Qy  228 LIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALME 287
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 liklicnvqameemmemkyntkkaplgkltvaqikagygslkagyysglkkiedciragqhgralme 300

Qy  288 ACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQ 347
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 acnefytriphdfglrtpplirtqkelsekiqllealgdieiaiklvktelqspehpldq 360
```

RESULT 13
AAB47031
ID   AAB47031 standard; Protein; 287 AA.
XX
AC   AAB47031;
XX
DT   29-MAR-2001 (first entry)
XX
DE   C-terminal fragment of hPARP2.
XX
KW   Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
KW   inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
KW   infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
KW   gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
KW   endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
KW   toxic shock syndrome; multiple organ injury syndrome; vasculitis;
KW   hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
KW   eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
KW   chronic obstructive pulmonary disease; silicosis; reperfusion injury;
KW   pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
KW   bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
KW   scar tissue formation; atherosclerosis; systemic lupus erythematosus;
KW   autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
KW   graft versus host disease; allograft rejection; cystic fibrosis;
KW   chronic glomerulonephritis; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
KW   inflammatory dermatosis; contact dermatitis; atopic dermatitis;
KW   psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
KW   Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
KW   hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
KW   leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;
KW   expressed sequence tag; EST; RACE; PCR; amplify; primer;
KW   polymerase chain reaction.
XX
OS   Synthetic.
XX
PN   WO200077179-A2.
XX
PD   21-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16629.
XX
PR   16-JUN-1999; 99US-0139543.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR   WPI; 2001-025335/03.
DR   N-PSDB; AAC85320.

KW   inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
KW   infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
KW   gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
KW   endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
KW   toxic shock syndrome; multiple organ injury syndrome; vasculitis;
KW   hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
KW   eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
KW   chronic obstructive pulmonary disease; silicosis; reperfusion injury;
KW   pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
KW   bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
KW   scar tissue formation; atherosclerosis; systemic lupus erythematosus;
KW   autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
KW   graft versus host disease; allograft rejection; cystic fibrosis;
KW   chronic glomerulonephritis; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
KW   inflammatory dermatosis; contact dermatitis; atopic dermatitis;
KW   psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
KW   Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
KW   hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
KW   leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;
KW   expressed sequence tag; EST; RACE; PCR; amplify; primer;
KW   polymerase chain reaction.
XX
OS   Synthetic.
XX
PN   WO200077179-A2.
XX
PD   21-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16629.
XX
PR   16-JUN-1999; 99US-0139543.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR   WPI; 2001-025335/03.
DR   N-PSDB; AAC85314.
XX
PT   New human poly(ADP-ribose) polymerase for treating inflammatory,
PT   neurological, cardiovascular, or neoplastic tissue growth disorders,
PT   such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
PT   metastasis -
XX
PS   Example 2: Page 100-01; 129pp; English.
XX
CC   This sequence is encoded by the 5' fragment of the human parp2
CC   (poly(ADP-ribose) polymerase) cDNA. The cDNA sequence was amplified
CC   using the primers given in AAC85307-13. DNA derived from spleen and
CC   testis cDNA libraries were used as templates. The primers given
CC   in AAC85315-17 were used to confirm the sequence of the 5'-hPARP2
CC   fragment. The amplified fragments P2-1 and P2-9 (AAC85318-19) were
CC   cloned to help determine that hPARP2 has an open reading frame of
CC   1080 nucleotides. The protein of the invention, hPARP2, causes
CC   the covalent addition of polymers of ADP-ribose to protein targets.
CC   hPARP2 activity is induced in many instances of oxidative stress or
CC   during inflammation where there is direct damage to the DNA.
CC   hPARP2 may be used to identify antagonists which may be used to
CC   treat a human having a disorder mediated by PARP2 activity, such as,
CC   inflammatory, neurological, cardiovascular, or neoplastic tissue
CC   growth disorders. hPARP2 and antibodies to it, can also be used
CC   to diagnose these conditions.
XX
SQ   Sequence 360 AA;

Query Match   59.7%; Score 1790.5; DB 22; Length 360;
Best Local Similarity   96.4%; Pred. No. 5.4e-161;
Matches 347; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

Qy   1 MAARRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRESKKMPVAGGKANKD 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
XX
PT   New human poly(ADP-ribose) polymerase for treating inflammatory,
PT   neurological, cardiovascular, or neoplastic tissue growth disorders,
PT   such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
PT   metastasis -
XX
PS   Example 2; Page 101-02; 129pp; English.
XX
CC   This sequence is encoded by the 3' fragment of the human parp2
CC   (poly(ADP-ribose) polymerase) cDNA. The cDNA sequence was amplified
CC   using the primers given in AAC85315-17.  DNA derived from a testis
CC   cDNA library was used as a template.  The protein of the invention,
CC   hPARP2, causes the covalent addition of polymers of ADP-ribose to
CC   protein targets. hPARP2 activity is induced in many instances of
CC   oxidative stress or during inflammation where there is direct
CC   damage to the DNA. hPARP2 may be used to identify antagonists
CC   which may be used to treat a human having a disorder mediated by
CC   PARP2 activity, such as, inflammatory, neurological, cardiovascular,
CC   or neoplastic tissue growth disorders. hPARP2 and antibodies to it,
CC   can also be used to diagnose these conditions.
XX
SQ   Sequence   287 AA;

Query Match            50.6%;   Score 1518;  DB 22;  Length 287;
Best Local Similarity 100.0%;   Pred. No. 2.8e-135;
Matches  287;    Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     284 ALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIETAIKLVKTELQSPEH 343
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 almeacnefytriphdfglrtpplirtqkelsekiqllealgdietaiklvktelqspeh 60

Qy     344 PLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAF 403
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 pldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlldlfevekdgekeaf 120

Qy     404 REDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANVCFA 463
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 redlhnrmllwhgsrmsnwvgilshglriappeapitgymfgkgiyfadmssksanvcfa 180

Qy     464 SRLKNTGLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTV 523
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 srlkntglllllsevalgqcnelleanpkaegllqgkhstkglgkmapssahfvtlngstv 240

Qy     524 PLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNFLQLW 570
               |||||||||||||||||||||||||||||||||||||||||||||||
Db     241 plgpasdtglnpdgytlnpdgytlnyneyivynpnqvrmryllkvqfnflqlw 287

RESULT 14
AAU20129
ID   AAU20129 standard; Protein; 294 AA.
XX
AC   AAU20129;
XX
DT   06-DEC-2001  (first entry)
XX
DE   Human DNA repair and processing polypeptide #14.
XX
KW   DNA processing; human; mouse; rabbit; goat; horse; cat; gene therapy;
KW   chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW   cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW   ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW   hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW   cerebrovascular disorder; nervous system disorder; bacterial infection;
KW   fungal infection; viral infection; ocular disorder; endocrine disorder;
KW   gastrointestinal disorder; renal disorder; respiratory disorder; dog;
KW   wound healing; skin aging; organ transplantation; tissue regeneration;
KW   anti-infertility; DNA repair protein.
XX
OS   Homo sapiens.
XX
```

```
PN   WO200155204-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01336.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
```

us-09-701-586b-2.rag

```
XX
DR    WPI; 2001-496846/54.
DR    N-PSDB; AAS32196.
XX
PT    Nucleic acids encoding human polypeptides, useful for preventing,
PT    diagnosing and/or treating e.g. cancers, Parkinson's disease and
PT    diabetic retinopathy -
XX
PS    Claim 11; SEQ ID No 56; 460pp; English.
XX
CC    Sequences AAU20116-AAU20147 represent the DNA repair and processing
CC    polypeptides of the invention. DNA repair and processing polypeptides and
CC    their associated polynucleotides are useful in the diagnosis, treatment
CC    and prevention of various types of disorders in e.g. humans, mice,
CC    rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological
CC    condition can be determined by detecting the presence or absence of a
CC    mutation in a DNA repair and processing polynucleotide. The treatable
CC    disorders include autoimmune diseases such as rheumatoid arthritis,
CC    hyperproliferative disorders such as neoplasms of the breast or liver,
CC    cardiovascular disorders such as cardiac arrest, cerebrovascular
CC    disorders such as cerebral ischaemia, nervous system disorders such as
CC    Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC    ocular disorders such as corneal infection, endocrine disorders such as
CC    premature labour and infertility, gastrointestinal disorders such as
CC    Crohn's disease, renal disorders such as glomerulonephritis and
CC    respiratory disorders such as asthma and pleurisy. The polypeptides can
CC    also be used to aid wound healing, to prevent skin aging due to sunburn,
CC    to maintain organs before transplantation, to regenerate tissues and in
CC    chemotaxis.
CC    Note: The sequence data for this patent did not form part of the printed
CC    specification, but was obtained in electronic format directly from WIPO
CC    at ftp.wipo.int/pub/published_pct_sequences.

Query Match             48.6%;    Score 1456;  DB 22;   Length 294;
Best Local Similarity   97.9%;    Pred. No. 2.2e-129;
Matches  280; Conservative    1; Mismatches    5; Indels    0; Gaps    0;

Qy    192 QMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEMMEMKYNTKK 251
          : |||||||||||||||||||||||||:||||||:|| ||||||||||||||||||||
Db      1 exdyatntqdeeetkkeeslksplkpesqldlrvqeliklicnvqamemmemkyntkk 60

Qy    252 APLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQ 311
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     61 apigkltvaqikagyqslkkiedciragghgxalmeacnefytriphdfglrtpplirtq 120

Qy    312 KELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQ 371
          |||:|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db    121 kelxekiqllealgdieiaiklvktelqspehpldqhyrnlhcalrpldhesyefkvisq 180

Qy    372 YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLHNRMLLWHGSRMSNWVGILSHGLR 431
          ||||| |||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ylqsthapthsdytmtlldlfevekdgekeafredlhnrmllwhgsrmsnwvgilshglr 240

Qy    432 IAPPEAPITGYMPGKGIYFADMSSKSANYCFASRLKNTGLLLLSEV 477
          ||||| |||| |||||||||||||||||||||||||||||||||||
Db    241 iappeaxitgymxgkgiyfadmssksanycfasrlkntglllsev 286


RESULT 15
AAU21810
ID    AAU21810 standard; Protein; 294 AA.
XX
AC    AAU21810;
XX
DT    06-DEC-2001  (first entry)
XX
DE    Novel human neoplastic disease associated polypeptide #243.
XX
KW    Human; neoplastic disease associated polypeptide; cancer;
KW    hyperproliferative disorder; neural disorder; immune system disorder;
KW    muscular disorder; reproductive disorder; gastrointestinal disorder;
KW    pulmonary disorder; cardiovascular disorder; renal disorder;
KW    neuroprotective; cytostatic; anti inflammatory; vasotropic.
```

```
PR    29-SEP-2000;    2000US-0236327.
PR    29-SEP-2000;    2000US-0236367.
PR    29-SEP-2000;    2000US-0236368.
PR    29-SEP-2000;    2000US-0236369.
PR    29-SEP-2000;    2000US-0236370.
PR    02-OCT-2000;    2000US-0236802.
PR    02-OCT-2000;    2000US-0237037.
PR    02-OCT-2000;    2000US-0237038.
PR    02-OCT-2000;    2000US-0237039.
PR    02-OCT-2000;    2000US-0237040.
PR    13-OCT-2000;    2000US-0239935.
PR    13-OCT-2000;    2000US-0239937.
PR    20-OCT-2000;    2000US-0240960.
PR    20-OCT-2000;    2000US-0241221.
PR    20-OCT-2000;    2000US-0241785.
PR    20-OCT-2000;    2000US-0241786.
PR    20-OCT-2000;    2000US-0241787.
PR    20-OCT-2000;    2000US-0241808.
PR    20-OCT-2000;    2000US-0241809.
PR    20-OCT-2000;    2000US-0241826.
PR    01-NOV-2000;    2000US-0244617.
PR    08-NOV-2000;    2000US-0246474.
PR    08-NOV-2000;    2000US-0246475.
PR    08-NOV-2000;    2000US-0246476.
PR    08-NOV-2000;    2000US-0246477.
PR    08-NOV-2000;    2000US-0246478.
PR    08-NOV-2000;    2000US-0246523.
PR    08-NOV-2000;    2000US-0246524.
PR    08-NOV-2000;    2000US-0246525.
PR    08-NOV-2000;    2000US-0246526.
PR    08-NOV-2000;    2000US-0246527.
PR    08-NOV-2000;    2000US-0246528.
PR    08-NOV-2000;    2000US-0246532.
PR    08-NOV-2000;    2000US-0246609.
PR    08-NOV-2000;    2000US-0246610.
PR    08-NOV-2000;    2000US-0246611.
PR    08-NOV-2000;    2000US-0246613.
PR    17-NOV-2000;    2000US-0249207.
PR    17-NOV-2000;    2000US-0249208.
PR    17-NOV-2000;    2000US-0249209.
PR    17-NOV-2000;    2000US-0249210.
PR    17-NOV-2000;    2000US-0249211.
PR    17-NOV-2000;    2000US-0249212.
PR    17-NOV-2000;    2000US-0249213.
PR    17-NOV-2000;    2000US-0249214.
PR    17-NOV-2000;    2000US-0249215.
PR    17-NOV-2000;    2000US-0249216.
PR    17-NOV-2000;    2000US-0249217.
PR    17-NOV-2000;    2000US-0249218.
PR    17-NOV-2000;    2000US-0249244.
PR    17-NOV-2000;    2000US-0249245.
PR    17-NOV-2000;    2000US-0249264.
PR    17-NOV-2000;    2000US-0249265.
PR    17-NOV-2000;    2000US-0249299.
PR    17-NOV-2000;    2000US-0249300.
PR    01-DEC-2000;    2000US-0250160.
PR    01-DEC-2000;    2000US-0250391.
PR    05-DEC-2000;    2000US-0251030.
PR    05-DEC-2000;    2000US-0251988.
PR    05-DEC-2000;    2000US-0256719.
PR    06-DEC-2000;    2000US-0251479.
PR    08-DEC-2000;    2000US-0251856.
PR    08-DEC-2000;    2000US-0251868.
PR    08-DEC-2000;    2000US-0251869.
PR    08-DEC-2000;    2000US-0251989.
PR    08-DEC-2000;    2000US-0251990.
PR    11-DEC-2000;    2000US-0254097.
PR    05-JAN-2001;    2001US-0259678.
XX
PA    (HUMA-) HUMAN GENOME SCI INC.
XX
PI    Rosen CA,  Barash SC,  Ruben SM;
```

```
XX
OS   Homo sapiens.
XX
PN   WO200155163-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01358.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   11-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   14-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   18-AUG-2000; 2000US-0225759.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.

PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
```

PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-465558/50.
DR   N-PSDB; AAS35009.
XX
PT   Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT   diagnose diseases or disorders associated with aberrant expression or
PT   activity of polypeptides, and for treating cancers, rheumatoid
PT   arthritis -
XX
PS   Claim 11; SEQ ID No 537; 687pp; English.
XX
CC   The present invention relates to the isolation of novel human neoplastic
CC   disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
CC   sequences encoding for these polypeptides. The sequences of the
CC   invention are useful in the diagnosis, treatment, prevention and/or
CC   prognosis of disorders involving neoplastic disease such as
CC   hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC   cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC   astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC   also be useful for treating other disorders such as neural disorders,
CC   immune system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC   and renal disorders. The polynucleotide sequences of the invention are
CC   also useful in gene therapy. AAU21568-AAU21851 represent the novel human
CC   neoplastic disease associated polypeptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   294 AA;

Query Match   48.6%;   Score 1456; DB 22; Length 294;
Best Local Similarity   97.9%;   Pred. No. 2.2e-129;
Matches   280; Conservative   1; Mismatches   5; Indels   0; Gaps   0;

Qy   192 QMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEMMMEMKYNTKK 251
            : |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 exdyatntqdeeetkkeeslksplkpesqldlrvqeliklicnvqameemmmemkyntkk 60

Qy   252 APLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQ 311
            ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   61 aplgkltvaqikagyqslkkiedciragqhgxalmeacnefytriphdfglrtppli̅rtq 120

Qy   312 KELSEKIQLLEALGDTEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQ 371
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 kelxekiqllealgdieiaikivktelqspehpldqhyrnlhcalrpldhesyefkvisq 180

Qy   372 YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLHNRMLLWHGSRMSNWVGILSHGLR 431
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ylqsthapthsdytmtlldlfevekdgekeafredlhnrmllwhgsrmsnwvgilshglr 240

Qy   432 IAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEV 477
            |||||| |||| |||||||||||||||||||||||||||||||||||
Db   241 iappeaxitgynmxgkglyfadmssksanycfasrlkntglllllsev 286

Search completed: August 29, 2002, 07:57:12
Job time: 289 sec

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        August 29, 2002, 07:53:23 ; Search time 46.79 Seconds
                                            (without alignments)
                                            297.555 Million cell updates/sec

Title:          US-09-701-586B-2
Perfect score:  2998
Sequence:       1 MAARRRRSTGGRARALNES.........PNQVRMRYLLKVQFNFLQLW 570

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters:      231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued_Patents_AA:*
              1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
              2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
              3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
              4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
              5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
              6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.

                              SUMMARIES

         %
Result        Query
NO.   Score Match Length DB  ID                    Description
----- ----- ----- ------ --  --                    -----------
 1   1012.5  33.8  1013  4   US-08-860-886-2       Sequence  2, Appli
 2   1007.5  33.6  1014  4   US-08-078-347A-3      Sequence  3, Appli
 3    135    4.5   1327  4   US-09-196-387-2       Sequence  3, Appli
 4   120.5   4.0    469  4   US-09-052-089A-1      Sequence  1, Appli
 5    119    4.0    163  2   US-08-044-618-6       Sequence  6, Appli
 6   116.5   3.9    469  2   US-08-968-751-2       Sequence  2, Appli
 7    115    3.8    906  2   US-08-609-230A-9      Sequence  9, Appli
 8   112.5   3.8    560  2   US-08-336-618-22      Sequence 22, Appli
 9   112.5   3.8   1312  1   US-08-687-080-51      Sequence 51, Appl
10    111    3.7   3218  1   US-08-764-100-27      Sequence 27, Appl
11   110.5   3.7    700  2   US-07-720-589-2       Sequence  2, Appli
12   110.5   3.7    700  2   US-07-785-190-2       Sequence  2, Appli
13   110.5   3.7    700  5   US-08-235-836C-66     Sequence 66, Appli
14   110.5   3.7    700  3   PCT-US92-05539-2      Sequence  2, Appli
15   110.5   3.7    907  3   US-08-990-140-4       Sequence  4, Appli
16   110.5   3.7    907  4   US-09-546-238-4       Sequence  4, Appli
17   110.5   3.7    940  1   US-08-810-712-7       Sequence  7, Appli
18   110.5   3.7   1312  2   US-08-599-126-148     Sequence 148, App
19    110    3.7    683  6   5210183-3             Patent No. 5210183
20   109.5   3.7    693  3   US-08-235-836C-68     Sequence 68, Appl
21    109    3.6    926  3   US-08-755-587-187     Sequence 187, App
22   108.5   3.6    807  4   US-09-081-345-2       Sequence  2, Appli
23    108    3.6    458  1   US-08-336-618-24      Sequence 24, Appli
24   107.5   3.6    619  1   US-08-465-746-2       Sequence  2, Appli
25   107.5   3.6    619  1   US-08-214-164-2       Sequence  2, Appli
26   107.5   3.6    619  3   US-08-467-852A-3      Sequence  3, Appli
27   107.5   3.6    619  2   US-08-246-636-2       Sequence  2, Appli
28   107.5   3.6    619  2   US-08-247-491A-3      Sequence  3, Appli
29   107.5   3.6    619  2   US-08-319-795-2       Sequence  2, Appli
30   107.5   3.6    619  2   US-08-468-985-2       Sequence  2, Appli
31   107.5   3.6    619  3   US-08-312-949-2       Sequence  2, Appli
32   107.5   3.6    648  1   US-08-072-070-2       Sequence  2, Appli
33   107.5   3.6    648  1   US-08-469-434-2       Sequence  2, Appli
34   107.5   3.6    648  2   US-08-214-222-2       Sequence  2, Appli
35   107.5   3.6    648  2   US-08-467-852A-2      Sequence  2, Appli
36   107.5   3.6    648  2   US-08-468-718-2       Sequence  2, Appli
37   107.5   3.6    648  2   US-08-247-491A-2      Sequence  2, Appli
38   107.5   3.6    648  3   US-08-446-201-3       Sequence  3, Appli
39   107.5   3.6    695  1   US-08-127-499A-23     Sequence 23, Appl
40   107.5   3.6    695  1   US-08-482-847-23      Sequence 23, Appl
41   106.5   3.6    441  4   US-08-630-915A-34     Sequence 34, Appl
42   106.5   3.6    700  4   US-08-235-836C-74     Sequence 74, Appl
43    106    3.5    470  4   US-09-052-089A-2      Sequence  2, Appli
44    106    3.5    693  4   US-08-235-836C-72     Sequence 72, Appl
45    106    3.5   2482  1   US-08-328-254-6       Sequence  6, Appli

                                               ALIGNMENTS

RESULT 1
US-08-860-886-2
; Sequence 2, Application US/08860886
; Patent No. 635009
; GENERAL INFORMATION:
;  APPLICANT: Burkle, Alexander
;  APPLICANT: zur Hausen, Harald
;  APPLICANT: Jan-Heiner, Kupper
;  TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE
;  TITLE OF INVENTION: IN GENE THERAPY
;  NUMBER OF SEQUENCES: 2
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Pennie & Edmonds, LLP
;  STREET: 1155 Avenue of the Americas
;  CITY: New York
;  STATE: NY
;  COUNTRY: USA
;  ZIP: 10036-2811
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Diskette
;  COMPUTER: IBM Compatible
;  OPERATING SYSTEM: Windows
;  SOFTWARE: FastSEQ for Windows Version 2.0b
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/860,886
;  FILING DATE: 03-OCT-1997
;  CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER:
;  FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Coruzzi, Laura A
;  REGISTRATION NUMBER: 30,742
;  REFERENCE/DOCKET NUMBER: 8484-0028-999
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: 650-493-4935
;  TELEFAX: 650-493-5556
;  TELEX: 66141 PENNIE
;  INFORMATION FOR SEQ ID NO:   2:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 1013 amino acids
;  TYPE: amino acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
;  FRAGMENT TYPE: internal
; US-08-860-886-2

         Query Match                    33.8%;  Score 1012.5;  DB 4;  Length 1013;
```

Left column

```
Best Local Similarity  40.5%;  Pred. No. 6.6e-89;
Matches 221;  Conservative 111;  Mismatches 185;  Indels 29;  Gaps 11;

Qy   29 APEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAK 88
            ||  :||  |  :::.|   :    ||:::.| |||.||||   :
Db  493 APRGKSGAALSKKSKGQV------------KEEGINKSEKRMK-LTLKGGAAVDPD-SGL 538

Qy   89 VGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHS 148
        |||  .|  :| | ||:   |||||| :  || :   |||||| :
Db  539 EHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSN 598

Qy  149 LVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKE 208
        |  :  | :|:.|   |  |  |  .| | .:|  |||  :    | |:   |||
Db  599 KLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYG---QDEEAVKK- 653

Qy  209 ESLKSPLKPESQLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAGYQS 268
        .:|:  ||:::|.|   |:|:|   :|:  :|   |||||| ||:||
Db  654 --LTVNPGTKSKLPKPVQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSI 711

Qy  269 LKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIE 328
        :|  |:.||   :|: ||:| ||| :  ||| |:|.
Db  712 LSEVQQAVSQGSSDSQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIE 771

Qy  329 IAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMT 387
        :|.||  :  :|:| |:|| || |  |:: ||:  |||
Db  772 VAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATTHSAYDLE 831

Qy  388 LLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFG 445
        ::|:.|:::.||  .:|  |||| ||||||||||||:||
Db  832 VIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFG 891

Qy  446 KGIYFADMSSKSSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGL 505
        ||||||| :| |  :|| || |:|   :: :||:  ||:   | : | ||||
Db  892 KGIYFADMVSKSANYHTSQGDPIGLILGEVALGNMYELKHAS-HISRLPKGKHSVKGL 950

Qy  506 GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPQVRMRYLLKVQFN 565
        ||  ::| .: |||| :: |||.: |  ::|:||||| | :
Db  951 GKTTPDPSANISLDGVDVPLGTGISSGV---NDTSLLYNEYIVYDIAQVNLKYLLLKFN 1007

Qy  566 F-LQLW 570
        |  ||
Db 1008 FKTSLW 1013
```

```
RESULT 2
US-09-078-347A-3
; Sequence 3, Application US/09078347A
; Patent No. 6132968
; GENERAL INFORMATION:
;   APPLICANT:  Le, Xiao-Chun
;   APPLICANT:  Weinfeld, Michael
;   APPLICANT:  Xing, James Z.
;   TITLE OF INVENTION:  Methods for Quantitating Low Level
;   TITLE OF INVENTION:  Modifications of Nucleotide Sequences
;   NUMBER OF SEQUENCES:  3
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Medlen & Carroll, LLP
;     STREET:  220 Montgomery Street, Suite 2200
;     CITY:  San Francisco
;     STATE:  California
;     COUNTRY:  USA
;     ZIP:  94104
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/078,347A
;     FILING DATE:  13-MAY-1998
;     CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Carroll, Peter G.
```

Right column

```
;   REGISTRATION NUMBER:  32,837
;   REFERENCE/DOCKET NUMBER: UALB-03283
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (415) 705-8410
;     TELEFAX:   (415) 397-8338
INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  1014 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  not relevant
;     TOPOLOGY:  unknown
;   MOLECULE TYPE:  protein
US-09-078-347A-3

Query Match      33.6%;  Score 1007.5;  DB 4;  Length 1014;
Best Local Similarity  40.5%;  Pred. No. 2e-88;
Matches 221;  Conservative 109;  Mismatches 187;  Indels 29;  Gaps 11;

Qy   29 APEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAK 88
            ||  :||  |  :::.|   :    ||:::.| |||.||||   :
Db  494 APRGKSGAALSKKSKGQV------------KEEGINKSEKRMK-LTLKGGAAVDPD-SGL 539

Qy   89 VGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHS 148
        |||  .|  :| | ||:   |||||| :  || :   |||||| :
Db  540 EHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSN 599

Qy  149 LVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKE 208
        |  :  | :|:.|   |  |  |  .| | .:|  |||  :    | |:   |||
Db  600 KLEQMPSKEDAIEQFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYG---QDEEAVKK- 654

Qy  209 ESLKSPLKPESQLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAGYQS 268
        .:|:  ||:::|.|   |:|:|   :|:  :|   |||||| ||:||
Db  655 --LTVNPGTKSKLPKPVQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSI 712

Qy  269 LKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIE 328
        :|  |:.||   :|: ||:| ||| :  ||| |:|.
Db  713 LSEVQQAVSQGSSDSQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIE 772

Qy  329 IAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMT 387
        :|.||  :  :|:| |:|| || |  |:: ||:  |||
Db  773 VAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATTHSAYDLE 832

Qy  388 LLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFG 445
        ::|:.|:::.||  .:|  |||| ||||||||||||:||
Db  833 VIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFG 892

Qy  446 KGIYFADMSSKSSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGL 505
        ||||||| :| |  :|| || |:|   :: :||:  ||:   | : | ||||
Db  893 KGIYFADMVSKSANYHTSQGDPIGLILGEVALGNMYELKHAS-HISRLPKGKHSVKGL 951

Qy  506 GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPQVRMRYLLKVQFN 565
        ||  ::| .: |||| :: |||.: |  ::|:||||| | :
Db  952 GKTTPDPSANISLDGVDVPLGTGISSGVIDT---SLLYNEYIVYDIAQVNLKYLLLKFN 1008

Qy  566 F-LQLW 570
        |  ||
Db 1009 FKTSLW 1014
```

```
RESULT 3
US-09-196-387-2
; Sequence 2, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
;   APPLICANT:  de Lange, Titia
;   APPLICANT:  Smith, Susan
;   TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;   TITLE OF INVENTION:  OF USE THEREOF
;   NUMBER OF SEQUENCES:  12
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Klauber & Jackson
;     STREET:  411 Hackensack Avenue, 4th Floor
```

```
; ..    CITY: Hackensack
; ..    STATE: New Jersey
; ..    COUNTRY: USA
; ..    ZIP: 07601
; .. COMPUTER READABLE FORM:
; ..    MEDIUM TYPE: Floppy disk
; ..    COMPUTER: IBM PC compatible
; ..    OPERATING SYSTEM: PC-DOS/MS-DOS
; ..    SOFTWARE: PatentIn Release #1.0, Version #1.30
; .. CURRENT APPLICATION DATA:
; ..    APPLICATION NUMBER: US/09/196,387
; ..    FILING DATE:
; ..    CLASSIFICATION:
; .. PRIOR APPLICATION DATA:
; ..    APPLICATION NUMBER: 09/095,225
; ..    FILING DATE: June 10, 1998
; .. ATTORNEY/AGENT INFORMATION:
; ..    NAME: Jackson Esq., David A.
; ..    REGISTRATION NUMBER: 26,742
; ..    REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; .. TELECOMMUNICATION INFORMATION:
; ..    TELEPHONE: 201-487-5800
; ..    TELEFAX: 201-343-1684
; ..    TELEX: 133521
; .. INFORMATION FOR SEQ ID NO: 2:
; ..    SEQUENCE CHARACTERISTICS:
; ..       LENGTH: 1327 amino acids
; ..       TYPE: amino acid
; ..       STRANDEDNESS: single
; ..       TOPOLOGY: linear
; ..    MOLECULE TYPE: protein
; ..    HYPOTHETICAL: NO
; US-09-196-387-2

Query Match        4.5%;   Score 135;   DB 4;   Length 1327;
Best Local Similarity  24.4%;   Pred. No. 0.001;
Matches   75;   Conservative   31;   Mismatches  104;   Indels  98;   Gaps  15;

Qy   312 KELSEKIQL-LEALGDI------EIAI------KLVK-----TELQSPEHPLDQHYRNL 352
         ::::|:::|:|::|:|::       ||:|    ||:|       :|  |: ::|
Db  1044 RDIFETEQITLLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNP----YLTF 1099

Qy   353 HCA------LRPLDHESYEFKVISQYLQSTHAPTHSD---------YTMTLLDLFEVEKD 397
         |:     |::||  || :||: :||                   :|:  :
Db  1100 HCVNQGTILLDLAPEDKEYQSVEEEMQST-IREHRDGGNAGGIFNRYNVIRIQKVVNKKL 1158

Qy   398 GEKEAFRE------DLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIY 449
         |: |:        :: |:|:|||  |||| |  ||||
Db  1159 RERFCHRQKEVSEENHHHNERMLFHGSPFIN--AIIHKGF---DERHAYIGGMFGAGIY 1213

Qy   450 FADMSSKSANYCFASRLKNTGL------------LLLSEVALGQCNELLEANPKAEGL 495
         |:.||||  | :       |          :|    |  |:|:|
Db  1214 FAENSSKSNQYYYGIG-GGTGCPTHKDRSCYICHRQMLFCRVTIGK----------- 1258

Qy   496 LQGKHSTKGLGRMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVR 555
         :|||::::||      :| |:|| |:|: | : |:|  | : ||::|::||
Db  1259 -------------SFLQFSTMKMAHAPPGHHSVIG--RPSVNGLAYAEYVIYRGEQAY 1301

Qy   556 MRYLLKVQ 563
         |||::
Db  1302 PEYLITYQ 1309

RESULT 4
US-09-052-089A-1
; Sequence 1, Application US/09052089A
; Patent No. 6346605
; GENERAL INFORMATION:
; ..    APPLICANT: Lee, Soo Y.
; ..               Choi, Yongwon
; ..    TITLE OF INVENTION: SIGNAL TRANSDUCER FOR THE TNF RECEPTOR SUPER
; ..                        FAMILY, AND USES THEREOF
```

```
; .. NUMBER OF SEQUENCES: 16
; .. CORRESPONDENCE ADDRESS:
; ..    ADDRESSEE: David A. Jackson, Esq.
; ..    STREET: 411 Hackensack Ave, Continental Plaza, 4th
; ..            Floor
; ..    CITY: Hackensack
; ..    STATE: New Jersey
; ..    COUNTRY: USA
; ..    ZIP: 07601
; .. COMPUTER READABLE FORM:
; ..    MEDIUM TYPE: Floppy disk
; ..    COMPUTER: IBM PC compatible
; ..    OPERATING SYSTEM: PC-DOS/MS-DOS
; ..    SOFTWARE: PatentIn Release #1.0, Version #1.30
; .. CURRENT APPLICATION DATA:
; ..    APPLICATION NUMBER: US/09/052,089A
; ..    FILING DATE: 31-Mar-1998
; ..    CLASSIFICATION: <Unknown>
; .. ATTORNEY/AGENT INFORMATION:
; ..    NAME: Jackson Esq., David A.
; ..    REGISTRATION NUMBER: 26,742
; ..    REFERENCE/DOCKET NUMBER: 600-1-198 CIP 1
; .. TELECOMMUNICATION INFORMATION:
; ..    TELEPHONE: 201-487-5800
; ..    TELEFAX: 201-343-1684
; .. INFORMATION FOR SEQ ID NO: 1:
; ..    SEQUENCE CHARACTERISTICS:
; ..       LENGTH: 469 amino acids
; ..       TYPE: amino acid
; ..       STRANDEDNESS: single
; ..       TOPOLOGY: linear
; ..    MOLECULE TYPE: protein
; ..    HYPOTHETICAL: NO
; ..    FRAGMENT TYPE: <Unknown>
; ..    ORIGINAL SOURCE:
; ..       ORGANISM: Homo sapiens
; ..    SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-052-089A-1

Query Match        4.0%;   Score 120.5;   DB 4;   Length 469;
Best Local Similarity  20.7%;   Pred. No. 0.0047;
Matches   71;   Conservative   61;   Mismatches 118;   Indels  93;   Gaps  15;

Qy    83 PECTAKVGKAHVYCEGNDVYDVMLNQTNL--QFNNNKYYLI-QLLEDDAQRNFSVWMR 137
         ||:|:||| |:     ::| : |: |:   :| |  ::|:: || |
Db    47 PQCRIQVGKRTII-NKLFFDLAQEEENVLDREFLKNELDNVRAQLSQKDKEKRDS---- 100

Qy   138 WGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE----KFEKVPGKYDML- 191
         |: : ::|      :|  |::    ::|                 || ::||
Db   101 ----------------------------QVIIDTLRDTLEERNATVVSLQQALGKAEMLCS 133

Qy   192 ---QMDYATNTQDEEETKKEES--LKSPLKPESQDLRVQELIKLICNVQAMEEMMMEM 245
             :|| ::| ||:::   ::||:|:|:|::      :|::
Db   134 TLKKQMKYLEQQQDETKQAEEAGRLRSKMKTMEQIELLLQS-----QLPEVEMIRDM 187

Qy   246 KYNTKKAPLGKLIVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTP 305
         |: || |     :::|::| ||:|:||:      : |::|:|  ||  ||:::::
Db   188 -------GVGQSAVEQLAVYCVSLKKEYENLKEAR-----KASGEVADKLRKDLFSSRS 234

Qy   306 PLIRTQKELSE-KIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESY 364
         |:::||:|: |:|:| |:| :|:  |:       |||:  |:||
Db   235 KLQTVYSELDQAKLELKSAQKDLQSADK---EIMSLKKKLTMLQETLN--LPPVASET- 287

Qy   365 EFFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDL 407
         | |:               ||      : :|:::
Db   288 ----VDRLVLESPAPV----------EVNLKLRRPSFRDDI 314

RESULT 5
US-08-044-618-6
; Sequence 6, Application US/08044618
; Patent No. 5449605
```

GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257,696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654.0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-044-618-6

Query Match   4.0%;   Score 119; DB 1; Length 163;
Best Local Similarity 28.3%; Pred. No. 0.0012;
Matches 41; Conservative 27; Mismatches 57; Indels 20; Gaps 6;

Qy   29 APEDSSPAKKTRRCQPQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAK 88
             || | || :: :|::|| | : :| |||| | ||| :
Db   15 APRGKSGAVLSKKSKGQV---------KEEGINKSEKRMK-LTLKGGAAVDPD-SGL 60

Qy   89 VGKAHYYCEGNDVYDVMLNQTNLQFNNKKYYLIQLLEDDAQRNFSVWMRWGRVGK-MGQH 147
             | |||:| | : :: : | || ::|||:| :|||| :| || :|| :|
Db   61 EHSAHVLEKGGKVFSATLSLVDVVKGTNSYYKLKLLKDDKESRHWIFKSWDRVGTVIGSN 120

Qy   148 SLVACSGNLNKAKEI--FQKKFLDK 170
             || :|| | :| |::|
Db   121 KL--EQMLSKEDTIEHFMKLYEEK 142

RESULT 6
US-08-968-751-2
; Sequence 2, Application US/08968751
; Patent No. 5948643
; GENERAL INFORMATION:
; APPLICANT: Rubinfeld, Bonnee
; APPLICANT: Polakis, Paul G.
; APPLICANT: Ligenfelter, Carol
; APPLICANT: Vuong, Terilyn T.
; TITLE OF INVENTION: MODULATORS OF BRCA1 ACTIVITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ONYX Pharmaceuticals, Inc.
; STREET: 3031 Research Drive
; CITY: Richmond

; STATE: CA
; COUNTRY: USA
; ZIP: 94806
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,751
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Giotta, Gregory
; REGISTRATION NUMBER: 32,028
; REFERENCE/DOCKET NUMBER: ONYX1024 GG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 262-8710
; TELEFAX: (510) 222-9758
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-968-751-2

Query Match   3.9%;   Score 116.5; DB 2; Length 469;
Best Local Similarity 20.7%; Pred. No. 0.011;
Matches 71; Conservative 60; Mismatches 119; Indels 93; Gaps 15;

Qy   83 PECTAKVGKAHVYCEGNDVYDVMLNQTNL--QFNNNKYYLI-QLLEDDAQRNFSVWMR 137
             |:| |||| || ::|: |:|::|| :|:: |: ||:|| : :|:|:||||: :
Db   47 PQCRIQVGKRTII--NKLFFDLAQEEENVLDAEFLKNELDNVRAQLSQKDKEKRDS---- 100

Qy   138 WGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE---KFEKVPGKYDML- 191
                                   |: |: :|||| : : || || ::| :|:
Db   101 ----------------------QVIIDTLRDTLEERNATVVSLQQALGKAEMLCS 133

Qy   192 ---QMDYATNTQDEEETKKEES--LKSPLKPESQLDLRVQELIKLICNVQAMEEMMMEM 245
            || |: ||: :|: :| : || : :|:: :|||| ::|:| :|
Db   134 TLKKQMKYLEQQQDETKQAQEARRLRSKMKTMEQIELLLQS----QRPEVEEMIRDM 187

Qy   246 KYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTP 305
              |||| :|| |||:|: | |: |||| | |:| :
Db   188 ------GVGQSAVEQLAVYCVSLKKEYENLKEAR------KASGEVADKLRKDLFSSRS 234

Qy   306 PLIRTQKELSE-KIQLLEALGDIEIAIKIVKTELQSPEHPLDQHYRNLHCALRPLDHESY 364
             || | | |: | |::|: ||| :| : :|| |: :: :|
Db   235 KLQTVVSELDQAKLELKSAQKDLQSADK----EIMSLKKKLTMLQETLN--LPPVASET- 287

Qy   365 EFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDL 407
             : |:   : :   : :| :|:: : | :::| :||: :||
Db   288 ----VDRLVLESPAPV--------------EVNLKLRRPSFRDDI 314

RESULT 7
US-08-609-230A-9
; Sequence 9, Application US/08609230A
; Patent No. 5866333
; GENERAL INFORMATION:
; APPLICANT: Innerarity, Thomas L.
; APPLICANT: Qian, Xiaobing
; APPLICANT: Yamanaka, Shinya
; TITLE OF INVENTION: Screening Methods to Detect mRNA Targets
; TITLE OF INVENTION: of Editing Enzymes
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California

```
;  COUNTRY: USA
;  ZIP: 94111-3834
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/609,230A
;  FILING DATE: 01-MAR-1996
;  CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Fitts, Renee A.
;  REGISTRATION NUMBER: 35,136
;  REFERENCE/DOCKET NUMBER: 02307U-068100US
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: 650-326-2400
;  TELEFAX: 650-326-2422
;  INFORMATION FOR SEQ ID NO: 9:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 906 amino acids
;  TYPE: amino acid
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
US-08-609-230A-9
```

```
Query Match           3.8%;   Score 115; DB 2; Length 906;
Best Local Similarity   19.2%;   Pred. No. 0.047;
Matches  123;   Conservative  87;   Mismatches 178;   Indels 254;   Gaps    31;

Qy     2 AARRRRSTGGGRARALNE--SKRVNN---GNTAPEDSS---PAKKTRRCQRQESKKMPV  52
           :|  :|  |||:|  ::|   |:   ::|
Db    10 ASRFSASSGGGSRGAPQHYPKTAGNSEFLGKTPGQNAQKWIPARSTRRDDNS------A  63

Qy    53 AGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAK-------VG-KAHVYCEGNDVYDV 104
           | |:    |||| :|:|     ||  |    ::|       ||  :|
Db    64 ANNSANEKERHDAIFRKVRGILNK----LTPEKFDKLCLQLLNVGVESKLILKG-----V 114

Qy   105 MLNQTNLQFNNNKY-----YLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKA 159
           :|: | |:  ||      :|:  |   |: |: ||
Db   115 ILLIVDKALEEPKYSSLYAQLCLRLAEDAP-NFDGPAAEGQPGQ-----------KQ 159

Qy   160 KEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPES 219
           |  ||||:|     |   ::|| :|  :||   |  |:    ::|| :||  |  :|
Db   160 STTFRRLLISKLQDEFENRTRNVDV---YD------------KRE---NPLLPEH 196

Qy   220 QLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAG 279
           :    :|:|:|  ||::|::|       |||    :
Db   197 E-EQRAIAKIKMLGNIKFIGE------------DLIH--   224

Qy   280 QHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQ 339
           |  :      :  |  | ||   :||  | | ::|| :|| :| ::|
Db   225 ---ESILHKC--------LGKL-----------IKTLLEKKRVQLQDMGEDLECLCQIMRTVGP 263

Qy   340 SPEHP------LDQHY-------RNLHCALRPLDHESYEF-----------KVI 369
           :|:|       ||:|        ||: |:|| ::|
Db   264 RLDHERAKSLMDQYFARMCSLMLSKELPARIRFLLQDTVELREHHWVPRKAFLDNGPKTI 323

Qy   370 SQ-------YLQSTHAPTHSDYTMT--LLDLFEVEKD--GEKEAFREDLHNRML 412
           |:        ::|:      |     :| |:|  ::|
Db   324 NQIRQDAVKDLGVFIPAPMAQGRSDFFLEGPFMPPRMKDRDPLGGLRDMFGQ------ 376

Qy   413 LWHGSRMSNWVGIL------------SHGLRIAPPEAPITGYMFGKGIYFADMS 454
            | |:|:|:
Db   377 -MPGSGIGTGPGVIQDTFSPTMGRHRSNQLFNGHGGHIMPPTQSQFGEMGGKFM----- 429

Qy   455 SKSANVCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHS-TKGLGKMAPSSA 513
           :   |:       :|      ::| |:   :|            |   | :|
Db   430 --------------KSQGLSQLYHNQSQGL-------- 445

Qy   514 HFVTLNGSTVPLGPA-SDTGILNPDGYTLNVNEYIVYNPNQV 554
           : :: :    |::|  :  |  |||
Db   446 -LSQLQGGQSKDMPPRFSKKGQLNADEISLRPAQSFLMNKNQV 486
```

```
RESULT 8
US-08-336-618-22
;  Sequence 22, Application US/08336618
;  Patent No. 5763590
;  GENERAL INFORMATION:
;  APPLICANT: Peattie, Debra A.
;  APPLICANT: Harding, Matthew W.
;  APPLICANT: Livingston, David J.
;  TITLE OF INVENTION: ISOLATION OF AN Mr 52,000 FK506 BINDING
;  TITLE OF INVENTION: PROTEIN AND MOLECULAR CLONING OF A CORRESPONDING HUMAN
;  TITLE OF INVENTION: cDNA
;  NUMBER OF SEQUENCES: 32
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.
;  STREET: Two Militia Drive
;  CITY: Lexington
;  STATE: Massachusetts
;  COUNTRY: U.S.A.
;  ZIP: 02173
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/336,618
;  FILING DATE: 09-NOV-1994
;  CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: 07/963,325
;  FILING DATE: 16-OCT-1992
;  APPLICATION NUMBER: US 07/777,752
;  FILING DATE: 11-OCT-1991
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: PCT/
;  FILING DATE: 09-OCT-1992
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Granahan, Patricia
;  REGISTRATION NUMBER: 32,227
;  REFERENCE/DOCKET NUMBER: VPI91-06A
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: 617-861-6240
;  TELEFAX: 617-861-9540
;  INFORMATION FOR SEQ ID NO: 22:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 560 amino acids
;  TYPE: amino acid
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
US-08-336-618-22
```

```
Query Match           3.8%;   Score 112.5; DB 1; Length 560;
Best Local Similarity   20.7%;   Pred. No. 0.037;
Matches   79;   Conservative  59;   Mismatches 132;   Indels 111;   Gaps    18;

Qy    59 KDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKY 118
           |:|:: |||  |  |||  || | |:|    ||  ||  |:||
Db   134 EDLTEEDGGIIRRIRLGVKAMQGPNDGAMV---EVALEGYH-KDRLFDQRELCF----- 184

Qy   119 YLIQLLEDDA--------QRNFSVWMRWGRVGKMGQHSLVACSGNL----------N 157
            :|:| |||            | ::|   |: :|| :|| |
Db   185 ---EVGEGESLDLPCAWRRPFSAWRK----------ESIPSCTSNLAMLLAVWGRRGSRSH 232

Qy   158 KAKEI--FQKKFLDKTKNNWE-----DREKFEKVPG-----KYDMLQ 192
           || :|      |:|   :||          | | ||       | :|
Db   233 RTAELRYEVRLKSFEKAKESWEMSSARSWSRATYVKERGTAYFKEGKYKQALLQYKKIVS 292

Qy   193 -MDYATNTQDEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAME-EMMMEMKYNT 249
           : ::| |  :      ::       |: |   ::|   :
```

```
Db  293 WLEYESSFSGEEMQKVH-----ALRLASHLNLAMCHL-KLQAFSAAIESCNKALELDSNN 346
Qy  250 KKAPLGK----LTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTP 305
            :   |       :    | |  |    ::  ::|  :|
Db  347 EKGLFRRGEAHLAVNDFDLARADFQKVLQLYPSNKAAKTQLAVCQQ----------- 392
Qy  306 PLIRTQKELSEKIQL---LEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDH 361
            |:|::|:|   || | |:|    | |  :||   :||    |:|
Db  393 --RTRRQLAREKKLYANMFERLAEEE---HKVKAEVAAGDHPTDAERKSLPRVWPPMDT 446
Qy  362 ESYEFKVISQYLQSTHAPTHS 382
            |:|:|    ||
Db  447 K------MQSLPTHPHPHS 460
```

RESULT 9
US-08-687-080-51
; Sequence 51, Application US/08687080
; Patent No. 5965427
;   GENERAL INFORMATION:
;   APPLICANT: Gregory Dolganov
;   TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
;   NUMBER OF SEQUENCES: 175
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dehlinger & Associates
;   STREET: 350 Cambridge Avenue, Suite 250
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94306
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/687,080
;   FILING DATE: 17-JUL-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/592,126
;   FILING DATE: 26-JAN-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sholtz, Charles K.
;   REGISTRATION NUMBER: 38,615
;   REFERENCE/DOCKET NUMBER: 4600-0111.30
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 324-0880
;   TELEFAX: (415) 324-0960
;   INFORMATION FOR SEQ ID NO: 51:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1312 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;   INDIVIDUAL ISOLATE: TRANS. OF RAD50 cDNA (SEQ. 54), NT.
;   INDIVIDUAL ISOLATE: 389 TO 4324
US-08-687-080-51

Query Match   3.8%; Score 112.5; DB 2; Length 1312;
Best Local Similarity   21.3%; Pred. No. 0.15;
Matches   97; Conservative   71; Mismatches 150; Indels 137; Gaps   24;

```
Qy   5 RRRSTGGGR------ARALNESKRVN----NGNT-------APEDSSPAK 37
           | |:     ::  | | |      | |         |    :
Db 433 RDKKTGLGRIIELKSEILSKKQNELKNVYELQQLEGSSDRILELDQELIKAERELSKAE 492
Qy  38 KTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKAL----------LLKGKAPVDPEC 85
```

```
Db  493 KNSNVETLKMEVISLQNEKADLDRTLRKLDQEMEQLNHHTTRTQMEMLYKDKADKD-EQ 551
             :      :  : :  ||   ||   ::   |:   |   :|  :||
Qy   86 TAKVGKAHVYCEGNDVDYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMG 145
Db  552 IRKIKSRH-----SDELTSLLG----YFPNKKQ-----LED-----WL----- 580
            |:  :|     :|   ||  |        |    |  |    |||   ||
Qy  146 QHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEET 205
Db  581 -HS------------KSKEINQ------------TRDRLAKL--NKELASSEQNKNHINNELK 616
              ::                 :    : :     | | | :   |
Qy  206 KKEESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAG 265
Db  617 RKEEQLSS--------------YEDKLFDVCGSQDFESDLDRLKEEIEKSSKQRAMLAGATAV 665
            :|||  | |     ::  |::| |  | | :
Qy  266 Y-QSLKKIED----------CIRAGQHGRALMEACNEFYTRIPHDFGLRTPP--LIRTQKEL 314
            |  | :::|         :    :   :::   ::  ::  |: |||  |||
Db  666 YSQFTQLTDENQSCCPVCQRVFQTEAELQEVISDLQSK-----LRLAPDKLKSTESEL 719
Qy  315 SEKIQLL-EALGDIEIAIKLVKTELQSPEHP-LDQHYRNLHCALRPLDHESYEFKVISQY 372
            |: :|    ::   :      :          :
Db  720 KKKEKRRDEMLGLVPMRQSII--DLKEKEIPELRNKIQNVNRDIQRLKNDIEE----QET 773
Qy  373 LQSTHAPTHSD----YTMTLLDLFEVE-KDGEKE 401
            |             |:  | :|:|  :|||| |:| | ||
Db  774 LLGTIMPEEESAKVCLTDVTIMERFQMELKDVERK 808
```

RESULT 10
US-08-764-100-27
; Sequence 27, Application US/08764100
; Patent No. 5773700
;   GENERAL INFORMATION:
;   APPLICANT: van Grinsven J., Martinus Q.
;   APPLICANT: De Haan, Petrus T.
;   APPLICANT: Gielen L., Johannes J.
;   APPLICANT: Peters, Dirk
;   APPLICANT: Goldbach, Robert W.
;   TITLE OF INVENTION: Improvements in or Relating to Organic
;   TITLE OF INVENTION: Compounds
;   NUMBER OF SEQUENCES: 27
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sandoz Agro, Inc
;   STREET: 975 California Avenue
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC Compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/764,100
;   FILING DATE: 06-DEC-1996
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/214,064
;   FILING DATE:
;   APPLICATION NUMBER: US 08/032,235
;   FILING DATE: 17-MAR-1993
;   APPLICATION NUMBER: GB 9206016.9
;   FILING DATE: 19-MAR-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: NO. 577370Iris, Allen E.
;   REGISTRATION NUMBER: 34,490
;   REFERENCE/DOCKET NUMBER: 137-1061
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 354-3592
;   TELEFAX: (415) 857-1125
;   INFORMATION FOR SEQ ID NO: 27:
;   SEQUENCE CHARACTERISTICS:

```
;       LENGTH: 3218 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: unknown
;       TOPOLOGY: unknown
;       MOLECULE TYPE: protein
; US-08-764-100-27

Query Match             3.7%;   Score 111;  Pred. No. 0.93;   DB 1;  Length 3218;
Best Local Similarity   21.2%;
Matches   90;  Conservative   46;  Mismatches 148;  Indels 140;  Gaps  17;

Qy    196 ATNTQDEEETKKEES--LKSPLKPESQLDLRVQELIKL----ICNVQAMEMMMEMKYNT 249
              |:|::    || :  |::|   :|       |   ::|  :|  :
Db   2534 ASNSERPHETYRMETGLLYSILEPHESERLEGLYLYSLELYSILEVALLEASPLEPRSE- 2592

Qy    250 KKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYT----RIPHDFGLRTP 305
              |:|                        :  |          : |:| :|:|
Db   2593 --------------------------RGLLEP-HELYSTHRVALPRLIYSLYSP 2618

Qy    306 PLIRTQKELSEKIQLLEALGDIEIAIKLV-------KTELQSPEHPLDQHY--RNLH 353
              :: - ||:: |||   | -|  :|       :||:|||:||:||
Db   2619 RILELESERSERVALSERL-ESERCYSLYSGLYCYSPHELECYSSERGLNGLYLEARGCY 2677

Qy    354 CALRPL----DHESYEFKVISQYLQSTHAPTHSD-------- 383
              |  ||    | ::|||:|:|:|:|| :||||
Db   2678 SALAALASERPHEILESERASPILE-THRPHESERALAARGLETHRMETLYSGLNCYSSE 2736

Qy    384 -------------------YTWTLLDLFEVEKDGEKAFREDLHNRM 411
                                 :: |||:|:|::|| ||::|:|:
Db   2737 RLESERTHRTYRGLNILEALAVALLYSYGLYALAASNLYSTYRASNLETHRMETPHEC 2796

Qy    412 LLWHGS-RMSNWVGILSHGLRIAPPEA---PITGYMFGKGIYFADMSSKSANYCFASRLK 467
              ::  |  :|::|:||| |:       :|||:|||||:|||:|||::||||||:||
Db   2797 YSTHRSERASNPRGLLYSGLNLYSMETILEILEGLPRGLGLYASPLYSSERTY--RSERV 2854

Qy    468 NTGLLLLSEVALGQCNE------------LLEANPK--AEGLLQGKHSTKGL--GKMAPSSAH 514
              |:| ||||:      :::|: |::::||   |::| |:|
Db   2855 ALGLALALEVALASPSERVALALAVALLEGLPRGLASNILEILEASPGLNASNASPGLNH 2914

Qy    515 FVTL 518
              |:::
Db   2915 ISAL 2918

RESULT 11
US-07-720-589-2
; Sequence 2, Application US/07720589
; Patent No. 5324630
; GENERAL INFORMATION:
;    APPLICANT: LeFebvre, Rance B.
;    APPLICANT: Perng, Guey-Chen
;    TITLE OF INVENTION: Methods and Compositions for Diagnosing
;    TITLE OF INVENTION: Lyme Disease
;    NUMBER OF SEQUENCES: 2
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: James M. Heslin
;       STREET: 379 Lytton Avenue
;       CITY: Palo Alto
;       STATE: California
;       COUNTRY: USA
;       ZIP: 94301
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: Patentin Release #1.0,  Version #1.25
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/07/720,589
;       FILING DATE: 19910628
;       CLASSIFICATION: 424
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Heslin, James M.
```

```
;       REGISTRATION NUMBER: 29,541
;       REFERENCE/DOCKET NUMBER: 2307U-330/UC 90-017-1
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 415-326-2400
;       TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 2:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 700 amino acids
;       TYPE: AMINO ACID
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
; US-07-720-589-2

Query Match             3.7%;   Score 110.5;  DB 1;  Length 700;
Best Local Similarity   20.0%;  Pred. No. 0.084;
Matches  117;  Conservative   90;  Mismatches 227;  Indels 151;  Gaps  26;

Qy     49 KMPVAGGKANKDRTEDK--QDESVKALLLKGKAPVD--PECTAKVGKAHVYCEGNDVYDV 104
              |:||         :|:| |: :|||   |  : : |::
Db    189 KKDILSGNIESDIDIDSLVTDKVVAALLSENEAGVNFARDITDIQGETHK--ADQDKIDI 246

Qy    105 MLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQ 164
              |: :|     ||:|         ::|
Db    247 ELDNIH-ESDSNITETIENLRD--------------------QLEKATDEEH 277

Qy    165 KKFL---DKTKNNWEDREK----FEKVPGKYDMLQ--MDYATNT------QDEEETKKE 208
              |||    :    :|:||    |::|||:|:|     |:|||       |:|||:||
Db    278 KKEIESQVDAKKKQKEELDKKAINLDKAQQKLDSAEDNLDVQRNTVRREKIQEDINEINKE 337

Qy    209 ESLKSP---LKPESQLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAG 265
              : |      :: ::|::       :: ::
Db    338 KNLPKPGDVSSPKVDKQIQIKE-----SLEDLQEQLKETGDENQKREIEK--QIEIKKS 389

Qy    266 YQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALG 325
              :: |||   :    ||:||:|:  :::
Db    390 DEKILLKSKD-DKASKDGKAL-DLDRELNSK----ASSKEKSKAKEEITKGKSQKSLG 441

Qy    326 DIEIAIKLVKTELQS-PE------------------HPLDQHYR--NLHCALRPLDHESY- 364
              : |: |::|:|:|  ||                  ||:|:|:|  |:|:|: :|:|
Db    442 DLNNDENLMMPEDQKLPEVKKLDSKKEFKPVSEVEKLDKIFKSNNNVGELSPLDKSSYKD 501

Qy    365 ---EFKVISQ--YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE------------ 405
                 |:|::|   :|::      :|:|:|||::|:|||:|||::
Db    502 IDSKEETVNKDVNLQKTFKPQVKDQVTSLNEDLLTMSIDSSSPVFLEVIDPITNLGTLQLI 561

Qy    406 DLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASR 465
              |:||:| | | ||::|:||||  |:|:|||:|||:|||||:|||:|||::|||||||
Db    562 DLNTGVSL---KESTQQGIQRYGI------------YEREKDLVVIKMDSGKAKLQILDK 606

Qy    466 LKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKG--LGKMAPSSAHFVTLNGSTV 523
              || |||||:|||:    : :|:||:| ||||:|||:|||  |||||||:|||||||:||
Db    607 LENLKVVSESNFEINKNSSLYVDSKMILVAVRDKSSNDWRLAKFSPK----------- 654

Qy    524 PLGPASDTGILNPDGYTLNYNEYIVYNPQVRMRYL-LKVQFNFL 567
              ||||:| ||:||:||:|||:|||:|:|||::|::| |:|:|:|:
Db    655 -----------NLDEFILSENKIMPFTSFSVRKNFIYLQDEFKSL 688

RESULT 12
US-08-785-190-2
; Sequence 2, Application US/08785190
; Patent No. 5977339
; GENERAL INFORMATION:
;    APPLICANT: LeFebvre, Rance B.
;    APPLICANT: Perng, Guey-Chen
;    TITLE OF INVENTION: Methods and Compositions for Diagnosing
;    TITLE OF INVENTION: Lyme Disease
;    NUMBER OF SEQUENCES: 2
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: James M. Heslin
;       STREET: 379 Lytton Avenue
;       CITY: Palo Alto
```

Left column

```
      STATE: California
      COUNTRY: USA
      ZIP: 94301
COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/785,190
      FILING DATE:
      CLASSIFICATION:
PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 07/720,589
      FILING DATE:
ATTORNEY/AGENT INFORMATION:
      NAME: Heslin, James M.
      REGISTRATION NUMBER: 29,541
      REFERENCE/DOCKET NUMBER: 2307U-330/UC 90-017-1
TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415-326-2400
      TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 700 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
US-08-785-190-2
```

```
Query Match           3.7%;  Score 110.5;  DB 2;  Length 700;
Best Local Similarity 20.0%;  Pred. No. 0.084;
Matches 117;  Conservative 90;  Mismatches 227;  Indels 151;  Gaps 26;

Qy   49 KMPVAGGKANKDRTEDK--QDESVKALLLKGKAPVD--PECTAKVGKAHVYCEGNDVYDV 104
           | |     :| |   ||  :| ||:    | |  |    |   | || :| |
Db  189 KKDILSGNIESDIDIDSLVTDKVVAALLSENEAGVNFARDITDIQGETHK--ADQDKIDI 246

Qy  105 MLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQ 164
           |:   |  ::| | |
Db  247 ELDNIH-ESDSNITETIENLRD------------------QLEKATDEEH 277

Qy  165 KKFL----DKTKNNWEDREK---FEKVPGKYDMLQ--MDYATNT-----QDEEETKKE 208
        || |    |   |            :|  | ::|     :|          | ::|
Db  278 KKEIESQVDAKKKQKEELDKKAINLDKAQQKLDSAEDNLDVQRNTVREKIQEDINEINKE 337

Qy  209 ESLKSP---LKPESQLDRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAG 265
         :|         |     ::::|       ||         :|        :|  ::|
Db  338 KNLPKPGDVSSPKVDKQLQIKE-----SLEDLQELKETGDENQKREIEK--QIEIKKS 389

Qy  266 YQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALG 325
                     :|  |    :    :: | :|      |      :||
Db  390 DEKLLKSKD-DKASKDGKAL-DLDRELNSK------ASSKEKSKAKEEEITKGKSQKSLG 441

Qy  326 DIEIAIKLVKTELQS-PE------------HPLDQHYR--NLHCALRPLDHESY-- 364
        |:|         || |               | :|     || :   ||:|
Db  442 DLNNDENLMMPEDQKLPEVKKLDSKKEFKPVSEVEKLDKIFKSNNNVGELSPLDKSSYKD 501

Qy  365 ---EFKVISQ--YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE------------ 405
           |:::|    |  :                   |    |
Db  502 IDSKBETVNRDVNLQKTKPQVKDQVTSLNEDLTMSIDSSSPVFLEVIDPITNLGTLQLI 561

Qy  406 DLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFCKGIYFADMSSKSANYCFASR 465
        ||:|    :      ::         ||       ||
Db  562 DLNTGVSL----KESTQQGIQRYGI------YEREKDLVVIKMDSGKAKLQILDK 606

Qy  466 LKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKG--LGKMAPSSAHFVTLNGSTV 523
        | ||  :| |           :| |         | ||
Db  607 LENLKVVSESNFEINKNSSLYVDSKMILVAVRDKDSSNDWRLAKFSPK---------- 654

Qy  524 PLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYL-LKVQFNFL 567
           | |   :|    |:  :|  || ::   :|  |   |: :| |
```

Right column

```
Db  655 ----------NLDEFILSENKIMPFTSFSVRKNFIYLQDEFKSL 688

RESULT 13
US-08-235-836C-66
; Sequence 66, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
;     APPLICANT: Dunn, John J.
;     APPLICANT: Luft, Benjamin J.
;     TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
;     TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
;     NUMBER OF SEQUENCES: 144
;     CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Brookhaven National Laboratory
;     STREET:
;     CITY: Upton
;     STATE: NY
;     COUNTRY: USA
;     ZIP: 11973
; COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/235,836C
;     FILING DATE: 29-APR-1994
;     CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/148,191
;     FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
;     NAME: Bogoslan, Margaret C.
;     REGISTRATION NUMBER: 25,324
;     REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (516) 282-7338
;     TELEFAX: (516) 282-3729
; INFORMATION FOR SEQ ID NO: 66:
;     SEQUENCE CHARACTERISTICS:
;     LENGTH: 700 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-235-836C-66
```

```
Query Match           3.7%;  Score 110.5;  DB 4;  Length 700;
Best Local Similarity 20.0%;  Pred. No. 0.084;
Matches 117;  Conservative 90;  Mismatches 227;  Indels 151;  Gaps 26;

Qy   49 KMPVAGGKANKDRTEDK--QDESVKALLLKGKAPVD--PECTAKVGKAHVYCEGNDVYDV 104
           | |     :| |   ||  :| ||:    | |  |    |   | || :| |
Db  189 KKDILSGNIESDIDIDSLVTDKVVAALLSENEAGVNFARDITDIQGETHK--ADQDKIDI 246

Qy  105 MLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQ 164
           |:   |  ::| | |
Db  247 ELDNIH-ESDSNITETIENLRD------------------QLEKATDEEH 277

Qy  165 KKFL----DKTKNNWEDREK---FEKVPGKYDMLQ--MDYATNT-----QDEEETKKE 208
        || |    |   |            :|  | ::|     :|          | ::|
Db  278 KKEIESQVDAKKKQKEELDKKAINLDKAQQKLDSAEDNLDVQRNTVREKIQEDINEINKE 337

Qy  209 ESLKSP---LKPESQLDRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAG 265
         :|         |     ::::|       ||         :|        :|  ::|
Db  338 KNLPKPGDVSSPKVDKQLQIKE-----SLEDLQELKETGDENQKREIEK--QIEIKKS 389

Qy  266 YQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALG 325
                     :|  |    :    :: | :|      |      :||
Db  390 DEKLLKSKD-DKASKDGKAL-DLDRELNSK------ASSKEKSKAKEEEITKGKSQKSLG 441

Qy  326 DIEIAIKLVKTELQS-PE------------HPLDQHYR--NLHCALRPLDHESY-- 364
```

Qy 165 KKFL----DKTKNNWEDREK----FEKVPGKYDMLQ-MDYATNT-----QDEEETKKE 208
Db 278 KKEIESQVDAKKKQKEELDKKAINLDKAQQKLDSAEDNLDVQRNTVREKIQEDINEINKE 337
Qy 209 ESLKSP---LKPESQLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAG 265
Db 338 KNLPKPGDVSSPKVDKQLQIKE-----SLEDLQEQLKETGDENQKREIEK--QEIKKS 389
Qy 266 YQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPLIRTQKELSEKIQLLEALG 325
Db 390 DEKLLKSKD-DKASKDGKAL-DLDRELNSK------ASSKEKSKAKEEITKGKSQKSLG 441
Qy 326 DIEIAIKLVKTELQS-PE------HPLDQHYR-NLHCALRPLDHESY- 364
Db 442 DLNNDENLMMPEDQKLPEVKKLDSKKEFKPVSEVEKLDKIFKSNNNVGELSPLDKSSYKD 501
Qy 365 --EFKVISQ--YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE-- 405
Db 502 IDSKEETVNKDVNLQKTKPQVKDQVTSLNEDLTTMSIDSSSPVFLEVIDPITNLGTLQLI 561
Qy 406 DLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASR 465
Db 562 DLNTGVSL---KESTQQGIQRYGI------YEREKDLVVIKMDSGKAKLQILDK 606
Qy 466 LKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKG--LGKMAPSSAHFVTLNGSTV 523
Db 607 LENLKVVSESNFEINKNSSLYVDSKMILVAVRDKDSSNDWRLAKFSPK--- 654
Qy 524 PLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYL-LKVQFNFL 567
Db 655 ---------NLDEFILSENKIMPFTSFSVRKNFIYLQDEFKSL 688

RESULT 15
US-08-990-140-4
; Sequence 4, Application US/08990140A
; Patent No. 6093795
; GENERAL INFORMATION:
; APPLICANT: Olsen, Henrik S.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Sonenberg, Nahum
; APPLICANT: Methot, Nathalie
; APPLICANT: Rom, Eran
; TITLE OF INVENTION: Human Prt1-like Subunit Protein (hPrt1) and Human
; TITLE OF INVENTION: eIF4GF-like Protein (p97) Genes
; FILE REFERENCE: 1488.0700001
; CURRENT APPLICATION NUMBER: US/08/990,140A
; CURRENT FILING DATE: 1997-12-12
; EARLIER APPLICATION NUMBER: US 60/033,151
; EARLIER FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-990-140-4

  Query Match   3.7%; Score 110.5; DB 3; Length 907;
  Best Local Similarity 20.5%; Pred. No. 0.13;
  Matches 127; Conservative 87; Mismatches 196; Indels 211; Gaps 33;

Qy   2 AARRRRSTGGGRARALNE-SKRVNN---GNTAPEDSS---PAKKTRRCQRQESKKMPV 52
Db   10 ASRFSASSGGGGSRGAPQHYPKTAGNSEFLGKTPGQNAQKWIPARSTRRDDNS-----A 63
Qy   53 AGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAK------VG-KAHVYCEGNDVYDV 104
Db   64 ANNSANEKERHDAIFRKVRGILNK----LTPEKFDKLCLELLNVGVESKLILKG-----V 114
Qy 105 MLNQTNLQFNNNKY----YLIQLLEDDAQRNFSVMRWGRVGKMGQHSLVACSGNLNKA 159

Db 442 DLNNDENLMMPEDQKLPEVKKLDSKKEFKPVSEVEKLDKIFKSNNNVGELSPLDKSSYKD 501
Qy 365 ---EFKVISQ--YLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE----- 405
Db 502 IDSKEETVNKDVNLQKTKPQVKDQVTSLNEDLTTMSIDSSSPVFLEVIDPITNLGTLQLI 561
Qy 406 DLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASR 465
Db 562 DLNTGVSL---KESTQQGIQRYGI------YEREKDLVVIKMDSGKAKLQILDK 606
Qy 466 LKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKG--LGKMAPSSAHFVTLNGSTV 523
Db 607 LENLKVVSESNFEINKNSSLYVDSKMILVAVRDKDSSNDWRLAKFSPK------ 654
Qy 524 PLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYL-LKVQFNFL 567
Db 655 ---------NLDEFILSENKIMPFTSFSVRKNFIYLQDEFKSL 688

RESULT 14
PCT-US92-05539-2
; Sequence 2, Application PC/TUS9205539
; GENERAL INFORMATION:
; APPLICANT: LeFebvre, Rance B.
; APPLICANT: Perng, Guey-Chen
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; TITLE OF INVENTION: Lyme Disease
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: James M. Heslin
;   STREET: 379 Lytton Avenue
;   CITY: Palo Alto
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94301
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US92/05539
;   FILING DATE: 19920629
;   CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
;   NAME: Heslin, James M.
;   REGISTRATION NUMBER: 29,541
;   REFERENCE/DOCKET NUMBER: 2307U-330/UC 90-017-1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-326-2400
;   TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 700 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
PCT-US92-05539-2

  Query Match   3.7%; Score 110.5; DB 5; Length 700;
  Best Local Similarity 20.0%; Pred. No. 0.084;
  Matches 117; Conservative 90; Mismatches 227; Indels 151; Gaps 26;

Qy   49 KMPVAGGKANKDRTEDK--QDESVKALLLKGKAPVD--PECTAKVGKAHVYCEGNDVYDV 104
Db 189 KKDILSGNIESDIDIDSLVTDKVVAALLSENEAGVNFARDITDIQGETHK--ADQDKIDI 246
Qy 105 MLNQTNLQFNNNKYYLIQLLEDDAQRNFSVMRWGRVGKMGQHSLVACSGNLNKAKEIFQ 164
Db 247 ELDNIH-ESDSNITETIENLRD----------QLEKATDEEH 277

```
Db   115  ILLIVDKALEEPKYSSLYAQLCLRLAEDAP-NFDGPAAEGQPGQ------------KQ  159

Qy   160  KEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPES  219
               !::: !:: :!:! :: |     ||       !:! ||
Db   160  STTFRRLLISKLQDEFENRTRNVDV---YD------------KRE---NPLLPEE  196

Qy   220  QLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAG  279
               :: : !!:: !:: :  |         !!!!           |!
Db   197  E-EQRAIAKIKMLGNIKFIGE--------------------DLIH--  224

Qy   280  QHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQ  339
                :::                        !:!       !|! : ::::|
Db   225  ---ESILHKC-------------IKTLLEKKRVQLKDMGEDLECLCQIMRTVGP  263

Qy   340  SPEHP----LDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEV  394
               :!|    :: : !:|    : | |  !:|
Db   264  RLDHERAKSLMDQYFARM-CSL------------MLSKELPARIRF--LLQDTVEL  304

Qy   395  EKDG--EKEAFREDLHNRMLLWHGSRMSNWV-------GI-----LSHGLR--------  431
                :!|        !:!:  :         |:   :: |:|
Db   305  REHHWVPRKAF---------LDNGPKTINQIRQDAVKDLGVFIPAPMAQGMRSDFFLEGP  355

Qy   432  IAPPE----------APITGYMFGKGI------YFADMSSKSANYCFASRLKNTGLLLLS  475
               ||         !  ! |!||||     : :         ::!|
Db   356  FMPPRMKMDRDPLGGLADMFGQMPGSGIGTGPGVIQDRFSPTMGRHRSNQLFNGHGGHIM  415

Qy   476  EVALGQCNELLEANPKAEGLLQGKHS-TKGLGKMAPSSAHFVLNGSTVPLGPA-SDTGI  533
                |:           !:!!| |: ::!|         |  : :: : |    |
Db   416  PPTQSQFGEMGGKFMKSQGLSQLYHNQSQGL--------LSQLQGQSKDMPPRFSKKGQ  466

Qy   534  LNPDGYTLNYNEYIVYNPNQV  554
               || :  || : : ! |||
Db   467  LNADEISLRPAQSFLMNKNQV  487
```

Search completed: August 29, 2002, 07:58:08
Job time: 285 sec

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 29, 2002, 07:57:23 ; Search time 36.19 Seconds
                                        (without alignments)
                                        609.841 Million cell updates/sec

Title:     US-09-701-586b-2
Perfect score: 2998
Sequence:  1 MAARRRRSTGGGRARALNES..........PNQVRMRYLLKVQFNFLQLW 570

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:  105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters:    105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2981.5	99.4	583	1	PPO2_HUMAN	Q9ugn5 homo sapien
2	2565.5	85.6	559	1	PPO2_MOUSE	O88554 mus musculu
3	1180	39.4	637	1	PPOL_ARATH	Q11207 arabidopsis
4	1027	34.3	1011	1	PPOL_CHICK	P26446 gallus gall
5	1022	34.1	1012	1	PPOL_CRIGR	Q9r152 cricetulus
6	1017	33.9	1012	1	PPOL_MOUSE	P11103 mus musculu
7	1016	33.9	1013	1	PPOL_RAT	P27008 rattus norv
8	1012.5	33.8	1013	1	PPOL_HUMAN	P09874 homo sapien
9	1009	33.7	1015	1	PPOL_BOVIN	P18493 bos taurus
10	984.5	32.8	998	1	PPOL_XENLA	P31669 xenopus lae
11	979.5	32.7	994	1	PPOL_DROME	P35875 drosophila
12	964	32.2	996	1	PPOL_SARPE	Q11208 sarcophaga
13	720	24.0	533	1	PPO3_HUMAN	Q9y6f1 homo sapien
14	526	17.5	538	1	YQN4_CAEEL	Q09525 caenorhabdi
15	331.5	11.1	135	1	PPOL_ONCMA	O08824 oncorhynchu
16	304.5	10.2	1724	1	PPOV_HUMAN	Q9ukk3 homo sapien
17	133	4.4	805	1	SEC6_YEAST	P32844 saccharomyc
18	130.5	4.4	1005	1	RA50_METJA	Q58718 methanococc
19	128	4.4	2230	1	GOG4_HUMAN	Q13439 homo sapien
20	126	4.2	880	1	RA50_PYRAB	Q9uzc8 pyrococcus
21	126	4.2	2058	1	MY10_HUMAN	Q9hd67 homo sapien
22	122	4.1	1938	1	MYHD_HUMAN	Q9ukx3 homo sapien
23	121.5	4.1	724	1	HMMR_HUMAN	O75330 homo sapien
24	121.5	4.1	794	1	HMMR_MOUSE	O00547 mus musculu
25	119.5	4.0	5430	1	ACF7_HUMAN	Q9upn3 homo sapien
26	118.5	4.0	1325	1	G160_MOUSE	P55937 mus musculu
27	117.5	3.9	1940	1	MYH3_CHICK	P02565 gallus gall
28	117	3.9	963	1	KINH_HUMAN	P33176 homo sapien
29	116.5	3.9	886	1	RA50_ARCFU	O29230 archaeoglob
30	116	3.9	504	1	YLF3_CAEEL	Q03573 caenorhabdi
31	115.5	3.9	540	1	YK26_YEAST	P36112 saccharomyc
32	115.5	3.9	600	1	PFR1_TRYBB	P22225 trypanosoma
33	115	3.8	1790	1	USO1_YEAST	P25386 saccharomyc
34	115	3.8	1846	1	MY5B_RAT	P70569 rattus norv
35	114.5	3.8	982	1	P115_MYCGE	P47540 mycoplasma
36	114.5	3.8	1959	1	MYH9_CHICK	P14105 gallus gall
37	114	3.8	734	1	YHU9_YEAST	P32900 saccharomyc
38	114	3.8	1230	1	SMC3_YEAST	P47037 saccharomyc
39	113	3.8	963	1	KINH_MOUSE	Q61768 mus musculu
40	112.5	3.8	1650	1	BP28_CAEEL	Q23495 caenorhabdi
41	112.5	3.7	1978	1	MYHB_CHICK	P10587 gallus gall
42	112	3.7	770	1	KLPA_EMENI	P28739 emericella
43	111.5	3.7	727	1	MFP1_ARATH	Q91w85 arabidopsis
44	111.5	3.7	2035	1	EVPL_MOUSE	Q9d952 mus musculu
45	111	3.7	876	1	PR02_YEAST	P20095 saccharomyc

ALIGNMENTS

```
RESULT 1
PPO2_HUMAN
ID   PPO2_HUMAN     STANDARD;      PRT;     583 AA.
AC   Q9UGN5; Q9Y6C8; Q9NUV2; Q9UMR4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE   (hPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoeger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE OF 2-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [3]
RP   SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fibroblast;
RX   MEDLINE=99268466; PubMed=10338144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shibashi T., Fujinori K.,
RA   Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA   Saito R., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA   Masuho Y., Kanehori K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC       SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor =
```

```
CC  -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC  -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC      produced by alternative splicing.
CC  -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN
CC      THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO
CC      DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;
CC      LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND
CC      THYMUS.
CC  -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC  --------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------------
DR  EMBL; AJ236912; CAB65088.1; -.
DR  EMBL; AF085734; AAD29857.1; ALT_INIT.
DR  EMBL; AJ236876; CAB41505.2; ALT_INIT.
DR  EMBL; AK001980; BAA92017.1; ALT_TERM.
DR  HSSP; P26446; 1A26.
DR  InterPro; IPR001290; PARP.
DR  InterPro; IPR004102; PARP_reg.
DR  Pfam; PF00644; PARP; 1.
DR  Pfam; PF02877; PARP_reg; 1.
KW  Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW  ADP-ribosylation; Alternative splicing.
FT  DNA_BIND      1     88       POTENTIAL.
FT  DOMAIN       86    583       NAD-BINDING (BY SIMILARITY).
FT  DOMAIN        4      7       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                               (POTENTIAL).
FT  DOMAIN       35     40       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                               (POTENTIAL).
FT  VARSPLIC     68     80       MISSING (IN ISOFORM 2).
FT  CONFLICT    447    447       P -> H (IN REF. 2).
FT  CONFLICT    481    481       N -> H (IN REF. 4).
SQ  SEQUENCE   583 AA;  66205 MW;  5B7AE8AE531836AF CRC64;
```

```
Query Match             99.4%;  Score 2981.5;  DB 1;  Length 583;
Best Local Similarity   97.8%;  Pred. No. 1.4e-187;
Matches  570;  Conservative    0;  Mismatches    0;  Indels   13;  Gaps    1;

Qy    1 MAARRRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAARRRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD 60

Qy   61 RTEDKQD------------ESVKALLLKGKAPVDPECTAKVGKAHVCEGNDYDVMLN 107
          |||||||            |||||||||||||||||||||||||||||||||||||||
Db   61 RTEDKQDGMPGRSWASKRVSESVKALLLKGKAPVDPECTAKVGKAHVCEGNDYDVMLN 120

Qy  108 QTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKF 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 QTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKF 180

Qy  168 LDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQE 227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQE 240

Qy  228 LIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALME 287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALME 300

Qy  288 ACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQ 347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQ 360

Qy  348 HYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDL 407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 HYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDL 420
```

```
Qy  408 HNRMLLWHGSRMSNWVGILSHGLRIAPPEPAPITGYMFGKGIYFADMSSKSANYCFASRLK 467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 HNRMLLWHGSRMSNWVGILSHGLRIAPPEPAPITGYMFGKGIYFADMSSKSANYCFASRLK 480

Qy  468 NTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGP 527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 NTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGP 540

Qy  528 ASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQOFNFLQLW 570
          ||||||||||||||||||||||||||||||||||||||||||||
Db  541 ASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQOFNFLQLW 583
```

```
RESULT  2
PPO2_MOUSE
ID  PPO2_MOUSE     STANDARD;      PRT;    559 AA.
AC  O08854; Q99N29;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  01-MAR-2002 (Rel. 41, Last annotation update)
DE  Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE  ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE  (mPARP-2).
GN  ADPRTL2 OR PARP2 OR ADPRT2.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC  TISSUE=Embryo;
RX  MEDLINE=99292755; PubMed=10364231;
RA  Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA  Muller S., Hoger T., Menissier-de Murcia J., de Murcia G.M.;
RT  "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)
RT  polymerase.";
RL  J. Biol. Chem. 274:17860-17868(1999).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=129Sv;
RX  MEDLINE=21179160; PubMed=11133988;
RA  Ame J.-C. Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,
RA  Niedergang C.P.;
RT  "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2
RT  (PARP-2) gene to the gene for RNase P RNA.";
RL  J. Biol. Chem. 276:11092-11099(2001).
RN  [3]
RP  SEQUENCE OF 9-559 FROM N.A.
RC  STRAIN=129/Sv X C57BL/6;
RX  MEDLINE=99268466; PubMed=10338144;
RA  Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT  "PADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT  related to truncated pADPRT homologues in plants and Caenorhabditis
RT  elegans.";
RL  FEBS Lett. 449:259-263(1999).
CC  -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC      SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.
CC  -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC      nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC  -!- SUBCELLULAR LOCATION: Nuclear.
CC  -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in
CC      testis followed by ovary.
CC  -!- INDUCTION: By high levels of DNA-damaging agents.
CC  -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC  --------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------------
```

DR   EMBL; AJ007780; CAA07679.1; -.
DR   EMBL; AF191547; AAK13253.1; -.
DR   EMBL; AF072521; AAC25415.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1341112; Adprt2.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND   1   65   POTENTIAL.
FT   DOMAIN   66   559   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   3   9   NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT   (POTENTIAL).
FT   DOMAIN   33   39   NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT   (POTENTIAL).
FT   CONFLICT   82   82   L -> V (IN REF. 2).
FT   CONFLICT   177   177   V -> I (IN REF. 2).
FT   CONFLICT   486   486   R -> Q (IN REF. 2).
SQ   SEQUENCE   559 AA; 63396 MW; E0AEDAEE412C1445 CRC64;

Query Match   85.6%; Score 2565.5; DB 1; Length 559;
Best Local Similarity   85.8%; Pred. No. 2e-160;
Matches 490; Conservative 39; Mismatches 29; Indels 13; Gaps 7;

Qy   1 MAARRRRSTGGGRARALNESKRVNNGNTAPEDSS-PAKKTRRCQRQESKKMPVAGGKANK 59
            ||:||| || | |||||| || || || || |: |||||
Db   1 MAPRRQRS-GSGR-RVLNEAKKVDNGNKATEDDSPPGKKMRTCQR----KGPMAGGK-DA 53

Qy   60 DRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYY 119
            |||||||::||||||||||||||||||||||||||||:||||||||||||||||||||
Db   54 DRTKDNRD-SVKTLLLKGKAPVDPECAAKLGKAHVYCEGDDVYDVMLNQTNLQFNNNKYY 112

Qy   120 LIQLLEDDAQRNFSVMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE 179
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113 LIQLLEDDAQRNFSVMRWGRVGKTGQHSLVTCSGDLNKAKEIFQKKFLDKTKNNWEDRE 172

Qy   180 KFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAME 239
            |||||||||||||||| :|||||||||||:|:::| |||||||||||||||||| |
Db   173 NFEKVPGKYDMLQMDYAASTQDESKTKEEET----LKPESQLDLRVQELLKLICNVQTME 228

Qy   240 EMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHD 299
            |||:||| || ||||||||||||||||||||||||||||||||| ||||||||||||
Db   229 EMMIEMKYDTKRAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALVEACNEFYTRIPHD 288

Qy   300 FGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPL 359
            |||:|| | :|:|| ||||||||||||||||||||||||:|| ||||||||||||||||
Db   289 FGLSIPPVVFREKELSDKVKLLEALGDIEIAIALKLVKSERQLEHPLDQHYRNLHCALRPL 348

Qy   360 DHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKKEAFREDLHNRMLLWHGSRM 419
            |||| |||||||||||||||:|| ||||||||:|||||||||||||| ||||||||||| :
Db   349 DHESNEFKVISQYLQSTHAPTHKDYTMTLLDVFEVEKEGEKKEAFREDLPNRMLLWHGSRL 408

Qy   420 SNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVAL 479
            |||||||||||: ||:||||||||||||||||||||||||||||||||||||||:|||||
Db   409 SNWVGILSHGLRVAPPEAPITGYMFGKGIYFADMSSKSANYCFASRIKNTGLLLLSEVAL 468

Qy   480 GQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGY 539
            |||||||||||| ||||:|||||||:|||||| || :|||||||||||||||||||| ||
Db   469 GQCNELLEANPKAQGLLRGKHSTKGMGKMAPSPAHFITLNGSTVPLGPASDTGILNPEGY 528

Qy   540 TLNYNEFIVYNPNQVRMRYLLKVQFNFLQLW 570
            ||||:||||:|||||||||||||:|||||||
Db   529 TLNYNEFIVYSPNQVRMRYLLKIQFNFLQLW 559

RESULT 3
PPOL_ARATH
ID   PPOL_ARATH   STANDARD;   PRT;   637 AA.
AC   Q11207;
DT   01-OCT-1996 (Rel. 34, Created)

DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   APP.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95269779; PubMed=7750552;
RA   Lepiniec L., Babiychuk E., Kushnir S., van Montagu M., Inze D.;
RT   "Characterization of an Arabidopsis thaliana cDNA homologue to animal
RT   poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 364:103-108(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   ---
DR   EMBL; Z48243; CAA88288.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND   1   140   POTENTIAL.
FT   DOMAIN   140   637   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   41   62   NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   637 AA; 72175 MW; 527A8F464605D127 CRC64;

Query Match   39.4%; Score 1180; DB 1; Length 637;
Best Local Similarity   43.8%; Pred. No. 8.2e-70;
Matches 252; Conservative 99; Mismatches 193; Indels 32; Gaps 10;

Qy   2 AARRRRSTGGGR---ARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKA 57
            : :|: :|| :   :||:| ::: : :|:   : :|
Db   81 AIKRGLDTTGTKKDLLERLCNDANNVSN------APVKSSNGTDEAEDD-------- 123

Qy   58 NKDRTEDKQDESVKALLLKGKAPVDPECTAKV-GKAHVYCEGNDVYDVMLNQTNLQFNNN 116
             ||:|||||||||||||||||||||:||||| ||||||||||||||||||||||||||
Db   124 NNGFEEEKKEEKIVTATKKGAAVLDQWIPDEIKSQYHVLQRGDDVYDAILNQTNVRDNNN 183

Qy   117 KYYILIQLLEDDAQRNFSVMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWE 176
             |||||||||||||||||||||||||||||||| |||| |||||||||| | |||||
Db   184 KFFVLQVLESDSKKTYMYTRWGRVGVGVKGQSKLDGPYDSWDRAIEIFTNKFNDKTKNYWS 243

Qy   177 DREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPE-SQLDLRVQELIKLICNV 235
             ||:|: :|||: :| :||| :::| ||| : ||| :::| || : ||||||||
Db   244 DREKFIPHPKSYTWLEMDYGKEENDSPVNNDIPSSSSEVKPEQSKLDTRVAKFISLICNV 303

```
Qy   236  QAMEEMMMEMKYNTKKAPLGKLIVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTR  295
               : :|||: || : ||::|: || :  :|| : ||  ||::|      |||||
Db   304  SMMAQHMEIGYNANKLPLGKISKSTISKGYEVLKRISEVIDRYDRTR-LEELSGEFYTV   362

Qy   296  TPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLH           353
          ||||| .: .:  :|:|  ||||::||||:|:| | : :|. | :|::|:
Db   363  IPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLLSVDPGLQDDPLIYHYQQLN 422

Qy   354  CALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE--DLHHNRM 411
          | |::| | :::|| ::|::|  |  ||  ||   ||  :| |   | ::|    |:| ||
Db   423  CGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVELAQLFRASRAVEADRFQQFSSSKNRM  482

Qy   412  LLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGL  471
          ||||||| ||| ||||:|||||||||| |||||||||:|||| |||||||: :||:|  :
Db   483  LLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKGVYFADMFSKSANYCYANTGANDGV  542

Qy   472  LLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTL-NGSTVPLGPASD  530
          |||.||||| .||||  : : : :| || |||| | :||..|::|:| ||: ||||| :
Db   543  LLLCEVALGDMNELLYSDYNADNLPPGKLSTKGVGCKTAPNPSEAQTLEDGVVVPLGKPVE 602

Qy   531  TGILNPDGVTLNEYIVYNPNQVRMRYLLKVQFNF                            566
             ||||||  |||||||| :|:|||||:::|:|
Db   603  RSCSKG---MLLYNEYIVNVEQIKMRYVIQVKFNY                           635
```

RESULT   4

PPOL_CHICK

```
ID   PPOL_CHICK     STANDARD;      PRT;   1011 AA.
AC   P26446;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=91340148; PubMed=1840535;
RA   Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;
RT   "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid
RT   sequence and comparison with mammalian enzyme sequences.";
RL   Gene 102:157-164(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.
RX   MEDLINE=98191351; PubMed=9521710;
RA   Ruf A., de Murcia G.M., Schulz G.E.;
RT   "Structure of the catalytic fragment of poly(AD-ribose) polymerase
RT   from chicken.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO
RP   895.
RX   MEDLINE=96355841; PubMed=8755499;
RA   Ruf A., de Murcia G.M., Schulz G.E.;
RT   "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived
RT   from crystal structures and homology modeling.";
RL   Biochemistry 37:3893-3900(1998).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.
RX   MEDLINE=98239716; PubMed=9571033;
RA   Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;
RT   "The mechanism of the elongation and branching reaction of poly(ADP-
RT   ribose) polymerase as derived from crystal structures and
RT   mutagenesis.";
RL   J. Mol. Biol. 278:57-65(1998).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC        PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC        ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC        CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC        TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC        EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC        nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC        AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC        FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC        THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC        AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```
```
DR   EMBL; X52690; CAA36917.1; -.
DR   PIR; JH0581; JH0581.
DR   PDB; 2PAW; 27-MAY-98.
DR   PDB; 1PAX; 15-MAY-97.
DR   PDB; 2PAX; 27-MAY-98.
DR   PDB; 3PAX; 27-MAY-98.
DR   PDB; 4PAX; 27-MAY-98.
DR   PDB; 1A26; 27-MAY-98.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; 3D-structure.
FT   DNA_BIND      1    370
FT   DOMAIN      371    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      382    458       BRCT.
FT   DOMAIN      523   1011       NAD-BINDING.
FT   ZN_FING      21     56       PARP-TYPE.
FT   ZN_FING     125    162       PARP-TYPE.
FT   DOMAIN      207    209       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     403    403       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     404    404       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     410    410       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     411    411       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     432    432       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     441    441       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     442    442       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     453    453       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     454    454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     468    468       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     481    481       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     509    509       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     510    510       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     517    517       ADP-RIBOSYL[N] (POTENTIAL).
```

```
FT   CONFLICT        895    895       A -> R (IN REF. 1).
SQ   SEQUENCE        1011 AA;  113520 MW;  261AED9383139144 CRC64;

Query Match
Best Local Similarity   40.8%;  Score 1027;  DB 1;  Length 1011;
Matches  218;  Conservative  110;  Mismatches  186;  Indels   20;  Gaps   11;

Qy    41  RCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGND  100
              :: |       | :::: ||  : ||  | :||| || | ||| :| |  ::|   :||
Db   494  KCSKPANMK---SAGKVKEEQGPSKSEKKMK-LTVKGGAAVDPD-SGLEDSAHVFEKGGK  548

Qy   101  VYDVMLNQTNLQFNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAK   160
              ::   |   ||   |||||||    :    | ||||||   :||    :|   ::
Db   549  IFSATLGLVDIVKGTNSYKLQLLEDDRESRYWVFRSWGRVGTVIGSNKLEQMPSKEDAV   608

Qy   161  EIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ  220
              ::| ||  |  :|| || ||   | | ||  :  |:   ::||||          |
Db   609  EHFLNLYEEKTGNSWHSK-NFTKYPKKFYPLEIDYG---QDEEAVRK---LTVSAGTKSK  661

Qy   221  LDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQ   280
              :|: | |||:| :     : |  ::  |:: |||||   :|   ||  : :|   :| |
Db   662  LAKPIQDLIKMIFDVESMKKAMVFEIDLQKMPLGKLSKRQIQSAYSILNEVQQAVSDGG   721

Qy   281  HGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVK-TELQ  339
              ||| |    ||||| ||||||:: ||| ::   |:| :|  ||  :||:|:  | ::
Db   722  SESQILDLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRGGNED  781

Qy   340  SPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE  399
              |  |:| :|: |  ||||| ||||: :|  ||::: ||| || | ::| : | |: :|
Db   782  GDKDPIDINYEKLRTDIKVVDKDSEEAKIIKQYVKNTHAATHNAYDLKVVEIFRIEREGE  841

Qy   400  KEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGVMFGKGIYFADMSSKS  457
              :|| |    ||||: ||||||: |: ||||:|||||||||||:||||||||||||| |||
Db   842  SQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGVMFGKGIYFADMVSKS  901

Qy   458  ANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVT  517
              ||||  :   : | |||| |||| |   |:  :|: |||:|||| |||||:|| ::: |
Db   902  ANYCHTSQADPIGLILLGEVALGNMYELKNASHITK-LPKGKHSVKGLGKTAPDPTATTT  960

Qy   518  LNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNF-LQLW        570
              |:|  ||||     || :  |  |||||||:|  :||  : |||: ||| | |||
Db   961  LDGVEVPLGNGISTGI---NDTCLLYNEYIVYDVAQVNLKYLLKLKFNYKTSLW       1011
```

RESULT 5
PPOL_CRIGR

```
ID   PPOL_CRIGR     STANDARD;      PRT;  1012 AA.
AC   Q9R152;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21276334; PubMed=11382339;
RA   Ganesh A., Phillips E., Thacker J., Meuth M.;
RT   "Suppression of the radiation-sensitive phenotype of hamster irs1 and
RT   irs2 strains selected for resistance to 3-aminobenzamide.";
RL   Int. J. Radiat. Biol. 77:609-616(2001).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC; CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; AF168781; AAD45817.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET       0      0
FT   DNA_BIND       1    371       BY SIMILARITY.
FT   DOMAIN       372    522       AUTOMODIFICATION DOMAIN (BY SIMILARITY).
FT   DOMAIN       386    475       BRCT.
FT   DOMAIN       523   1012       NAD-BINDING (BY SIMILARITY).
FT   ZN_FING       20     55       PARP-TYPE.
FT   ZN_FING      124    161       PARP-TYPE.
FT   DOMAIN       206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                 (POTENTIAL).
FT   DOMAIN       220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                 (POTENTIAL).
FT   MOD_RES      406    406       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      412    412       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      443    443       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      455    455       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      470    470       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      483    483       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      487    487       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      490    490       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      511    511       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES      518    518       ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE    1012 AA;  112400 MW;  8D617C4DBF0CB0F7 CRC64;

Query Match   34.1%;  Score 1022;  DB 1;  Length 1012;
Best Local Similarity   41.0%;  Pred. No. 3.2e-59;
Matches  227;  Conservative  107;  Mismatches  175;  Indels   44;  Gaps   15;

Qy    29  APEDSS--PAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECT   86
              || ::: |       | :::| | :::: ||  : ||  | :||| || | |||
Db   493  APKGKGKSAAPSKKSKGLYKEEG--------VNKSEKRMK-LTLKGGAAVDPD-S      535
```

```
Qy     87  AKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYLIQLLEDDAQRNFSVWMRWGRVGKMGQ  146
               ||  |  || :||:::  || |:  ||||||||| : : :  :|||||             |||||
Db    536  GLEHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVIG  595

Qy    147  HSLVACSSNLNRAKEIFQKKFLDKTKNNWEDREKFEKVPGKXYDMLQMDYATNTQDEEETK  206
               :. :::|||  :| ||   | :: ||| |  :: ||||                  |||||
Db    596  SNKLEQMPSKEDAVEHFMKLYEKTGNAWHSK-NFTKYPKKFYPLEIDYG---QDEEAVK  651

Qy    207  KEESLKSPLKP--ESQDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKA  264
                 |  ||  ||||| ||  |:|:  : :|:  :: :|  || |||| |:||
Db    652  ----KLTVKPGTKSKLPKAVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSKRQIQA  706

Qy    265  GYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPLIRTQKELSEKIQLLEAL  324
               |   |:|||:|||      :|| || :  ||||  |:|  :::|: ||||| |: :
Db    707  AYSILSEVQQAVSQGSSDSQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNL  766

Qy    325  GDIEIAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSD  383
               ||||| |: ||: : |: :| :::::  || :| ||||:|| :| :::|||||   |||
Db    767  LDIEVAYSLLRGGSDDSSKDPIDVNYELKTDIKVVDRDSEEAEVIRKYVKNTHATTHNA  826

Qy    384  YTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITG  441
               |  : :: ::||:|  ::  ::   ||| ::|| ||||  : |||  |||||||||||
Db    827  YDLEVMDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTG  886

Qy    442  YMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHS  501
               |||||||||||:||||||| ||:|:||| :|||:||:|||  | |||||  |:||:|||
Db    887  YMFGKGIYFADMVSKSANYCHTSQGDPIGLILLGEVALGNMYELKHASHISK-LPKGKHS  945

Qy    502  TKGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYT---LNYNEYIVYNPQVRMRY  558
               | |||:   :||| || ||| |  |  |:||| :||       |||:|| | :|::||
Db    946  VKGLGKTTPDPSASITLEGVEVPLG----TGI--PSGVNDTCLLYNEYIVYDIAQVNLKY  999

Qy    559  LLKVQFNF-LQLW  570
               |||:||:|| ||
Db   1000  LLKLKFNFKTSLW 1012
```

RESULT   6
PPOL_MOUSE
ID   PPOL_MOUSE   STANDARD;   PRT;   1012 AA.
AC   P11103; Q9JLX4; Q9QVQ3;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1) (msPARP)).
GN   ADPRT OR ADPRT1 OR ADPRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=BXSB;
RX   MEDLINE=89263780; PubMed=2498841;
RA   Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;
RT   "Sequence and organization of the mouse poly (ADP-ribose) polymerase
RT   gene.";
RL   Nucleic Acids Res. 17:3387-3401(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC   STRAIN=129/Sv X C57BL/6; TISSUE=Fibroblast;
RX   MEDLINE=20270268; PubMed=10809783;
RA   Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;
RT   "Characterization of sPARP-1. An alternative product of PARP-1 gene
RT   with poly(ADP-ribose) polymerase activity independent of DNA strand
RT   breaks.";
RL   J. Biol. Chem. 275:15504-15511(2000).
RN   [3]
RP   KNOCK-OUT.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jaeger S., Berghammer H.,

RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC   short form/sPARP-1; may be produced by alternative initiation.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X14206; CAA32421.1; -.
DR   EMBL; AF126717; AAF61293.1; ALT_INIT.
DR   PIR; S04200; S04200.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1340806; Adprt1.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP.reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative initiation.
FT   INIT_MET   0   0   BY SIMILARITY.
FT   CHAIN   1   1012   POLY [ADP-RIBOSE] POLYMERASE-1, LONG
FT   ISOFORM.
FT   CHAIN   521   1012   POLY [ADP-RIBOSE] POLYMERASE-1, SHORT
FT   ISOFORM.
FT   INIT_MET   521   521   FOR SHORT ISOFORM.
FT   DNA_BIND   1   371
FT   DOMAIN   372   522   AUTOMODIFICATION DOMAIN.
FT   DOMAIN   384   460   BRCT.
FT   DOMAIN   523   1012   NAD-BINDING.
FT   ZN_FING   20   55   PARP-TYPE.
FT   ZN_FING   124   161   PARP-TYPE.
FT   DOMAIN   206   208   NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN   220   225   NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES   406   406   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   412   412   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   434   434   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   436   436   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   443   443   ADP-RIBOSYL[N] (POTENTIAL).

```
FT   MOD_RES    444   444   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    447   447   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    455   455   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    483   483   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    487   487   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    490   490   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    511   511   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    512   512   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES    518   518   ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT   590   590   L -> V (IN REF. 2).
FT   CONFLICT   607   607   E -> D (IN REF. 2).
FT   CONFLICT   611   611   Q -> H (IN REF. 2).
FT   CONFLICT   628   628   N -> D (IN REF. 2).
FT   CONFLICT   678   678   D -> E (IN REF. 2).
FT   CONFLICT   702   702   R -> F (IN REF. 3).
FT   CONFLICT   716   716   Q -> E (IN REF. 2).
FT   CONFLICT   757   757   Q -> L (IN REF. 2).
FT   CONFLICT   856   856   R -> F (IN REF. 3).
FT   CONFLICT   981   981   A -> C (IN REF. 2).
SQ   SEQUENCE   1012 AA;  112968 MW;  4354C3E5F01B9439 CRC64;

Query Match       33.9%;  Score 1017;   DB 1;   Length 1012;
Best Local Similarity  40.1%;  Pred. No. 6.9e-59;
Matches  219;  Conservative 113;  Mismatches 178;  Indels 36;  Gaps 12;

Qy   31 EDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLKGKAPVDPECTAKVG 90
           ::::||:::||:::||::||
Db  497 KSAAPSKKSKGCFKEEG------------------VNKSEKRMK-LTLKGGAAVDPD-SGLEH 539

Qy   91 KAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLV 150
          ||| :| ::  :  :|| : |||||||||  :: ::     :||||:| |
Db  540 SAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRLGTVIGSNKL 599

Qy  151 ACSGNLNKAKEIFQKKFLDKTKNNWEDREKPEKVPGKYDMLQMDYATNTQDEEETKKEES 210
              :|| :||||  :   :: : || :: :||      :||||:
Db  600 EQMPSKEAVEQFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYG---QDEEAVK---- 651

Qy  211 LKSPLKP--ESQLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAGYQS 268
          :| || :  ||||| ::::| :: ::  :|| |   ::|||||::|||:||:|
Db  652 -KLTVKPGTKSKLPKPVQELVGMIFDVDSMKKALVEYEIDLQKMPLGKLSRRQIQAAYSI 710

Qy  269 LKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRRTPPLIRTQKELSEKIQLLEALGDIE 328
          | :::  :  :    :::||||:| |||||||:::||| :::  :|: ::  |||| |||
Db  711 LSEVQQPVSQGSSESQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIE 770

Qy  329 IAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMT 387
          :| ::|: :  :: : ||: ||  ::::|  :||:||| : |:||  :||:||::  ::
Db  771 VAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEVYIRKYVKNTHATTHNAYDLE 830

Qy  388 LLDLFEVKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFG 445
          ::||:||:::| :||:  ||||||:||||||:::| :|||:|||||||||||:|||||||
Db  831 VIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFG 890

Qy  446 KGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGL 505
          ||||||||:|||||| ::: :  :||:::|||||||  : :::|| :::||:||||:|||
Db  891 KGIYFADMVSKSANYCHTSQGDPIGLIMLGEVALGNMYELKHASHISK-LPKGKHSVKGL 949

Qy  506 GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFN 565
          || :   || :::|| ::|||  :: ||::|:| |||:|||| :: |||: :|||| :||
Db  950 GKTTPDPSASITLEGVEVPLGTGIPSGV---NDTALLYNEYIYDIAQVNLKYLLKLKFN 1006

Qy  566 F-LQLW 570
          | |:||
Db 1007 FKTSLW 1012
```

RESULT 7
PPOL_RAT

```
ID   PPOL_RAT       STANDARD;       PRT;   1013 AA.
AC   P27008; O35937;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Monocytes;
RX   MEDLINE=98046546; PubMed=9385436;
RA   Beneke S., Meyer R., Buerkle A.;
RT   "Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly
RT   (ADP-ribose) polymerase.";
RL   Biochem. Mol. Biol. Int. 43:755-761(1997).
RN   [2]
RP   REVISION TO 811.
RA   Beneke S., Meyer R., Buerkle A.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-11 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=92290013; PubMed=1601134;
RA   Potvin F., Thibodeau J., Kirkland J.B., Dandenault B.,
RA   Duchaine C., Poirier G.G.;
RT   "Structural analysis of the putative regulatory region of the rat
RT   gene encoding poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 302:269-273(1992).
RN   [4]
RP   SEQUENCE OF 514-1013 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=90027702; PubMed=2508731;
RA   Thibodeau J., Gradwohl G., Dumas C., Clairoux-Moreau S., Brunet G.;
RT   "Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase
RT   catalytic domain and analysis of mRNA levels during the cell cycle.";
RL   Biochem. Cell Biol. 67:653-660(1989).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1) + [ADP-D-ribosyl](N)-acceptor -
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------
DR   EMBL; U94340; AAC53544.1; -.
DR   EMBL; X65496; CAA46477.1; -.
DR   EMBL; X65497; CAA46478.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
```

```
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0
FT   DNA_BIND      1    372       BY SIMILARITY.
FT   DOMAIN      385    461       BRCT.
FT   DOMAIN      373    523       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      524   1013       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       PARP-TYPE.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES       1      1       BLOCKED (BY SIMILARITY).
FT   MOD_RES     407    407       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     413    413       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     435    435       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     437    437       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     456    456       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     484    484       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     491    491       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     513    513       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     519    519       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    638    638       Y -> H (IN REF. 4).
FT   CONFLICT    641    641       E -> A (IN REF. 4).
FT   CONFLICT    752    752       N -> D (IN REF. 4).
SQ   SEQUENCE   1013 AA;  112529 MW;  AA566F2B29BE97C0 CRC64;

Query Match             33.9%;
Best Local Similarity   40.8%;
Matches   224; Conservative  109; Mismatches  174; Indels  42; Gaps  14;

Qy    31 EDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVG 90
           ::::||:||: :|                      :: | | |||| ||:| | :
Db   498 KSAAPSKKSKSKGAVKEEG--------VNKSEKRMK-LTLKGGAAVDPD-SGLEH 540

Qy    91 KAHVCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLV 150
           ||| :| ||:|||| ||||| : |  ||| : ||| : |     ::|||||||:
Db   541 SAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLESDKESRSYWIFRSWGRVGTVIGSNKL 600

Qy   151 ACSGNLNKAKEIFQKKFLDKTKNMWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEES 210
           | :| ||||: :::|          :: |  |:|   |:  |||:|  |:|: :
Db   601 EQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYG---QDEEAVK---- 652

Qy   211 LKSPLKP--ESQLDLRVQELIKLICNVQAMEMMEMKYNTKKAPLGKLTVAQIKAGYQS 268
           |:|| ||   :| | |||| :|:   ::| | :||: |  |:||:| ::| |  :|
Db   653 -KLAVKPGTKSKLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQAAYSI 711

Qy   269 LKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIE 328
           | |||:||      : ::     :|||:|||||::| ||:  | :  ||::|||:
Db   712 LSEVQQAVSOGSSESOILDLSNRFYTLIPHDFGMKKPPLLNNTDSVQAKVEMLDNLLDIE 771

Qy   329 IAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISOQYLQSTHAPTHSDYTMT 387
           :| |:      | : :|:|:  |:|  :|: |:|::|||:|:   |  :|||:|| |:
Db   772 VAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNAYDLE 831

Qy   388 LLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFG 445
           :|:||:||  : ::|   |:||||:|||||||| :|  ||||||||||||||||:||||||
Db   832 VIDIFKIEREGESORYKPFROLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFG 891

Qy   446 KGIYFADMSSKSANYCFASRLKNTGLLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGL 505
           ||||||| |||:||||  |      | |||||||||| :|| :  : :  :|:||||:|||
Db   892 KGIYFADMVSKSANYCHTSOGDPIGLILLGEVALGNMYELKHASHISK-LPKGKHSVKGL 950

Qy   506 GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYT---LNYNEYIVYNPNQVRMRYLLKV 562
           || :| ||| |||||||:|||||| |:||  |||    |||||||:|| || || ||||
Db   951 GKTAPDPSASITLDGVEVPLG---TGI--PSGVNDTCLLYNEYIVDIAQVNLKYLLKL 1004

Qy   563 QFNF-LQLW 570
           ||| ||
Db  1005 KFNFKTSLW 1013

RESULT 8
PPOL_HUMAN
ID   PPOL_HUMAN    STANDARD;    PRT;   1013 AA.
AC   P09874;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT OR PPOL OR PARP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90091744; PubMed=2513174;
RA   Auer B., Nagl U., Herzog H., Schneider R., Schweiger M.;
RT   "Human nuclear NAD+ ADP-ribosyltransferase(polymerizing):
RT   organization of the gene.";
RL   DNA 8:575-580(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88076933; PubMed=3120710;
RA   Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S.,
RA   Suzuki H., Nyunoya H., Miwa M., Sugimura T.;
RT   "Nucleotide sequence of a full-length cDNA for human fibroblast
RT   poly(ADP-ribose) polymerase.";
RL   Biochem. Biophys. Res. Commun. 148:617-622(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88058958; PubMed=2824474;
RA   Kurosaki T., Ushiro H., Mitsuuchi Y., Suzuki S., Matsuda M.,
RA   Matsuda Y., Katunuma N., Kangawa K., Matsuo H., Hirose T.,
RA   Inayama S., Shizuta Y.;
RT   "Primary structure of human poly(ADP-ribose) synthetase as deduced
RT   from cDNA sequence.";
RL   J. Biol. Chem. 262:15990-15997(1987).
RN   [4]
RP   SEQUENCE OF 440-1013 FROM N.A.
RX   MEDLINE=88068596; PubMed=2891139;
RA   Cherney B.W., McBride O.W., Chen D., Alkhatib H., Bhatia K.,
RA   Hensley P., Smulson M.E.;
RT   "cDNA sequence, protein structure, and chromosomal location of the
RT   human gene for poly(ADP-ribose) polymerase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:8370-8374(1987).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87298455; PubMed=3113420;
RA   Suzuki H., Uchida K., Shina H., Sato T., Okamoto T., Kimura T.,
RA   Miwa M.;
RT   "Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and
RT   expression of its gene during HL-60 cell differentiation.";
RL   Biochem. Biophys. Res. Commun. 146:403-409(1987).
RN   [6]
RP   ERRATUM.
RA   Suzuki H., Uchida K., Shina H., Sato T., Okamoto T., Kimura T.,
RA   Miwa M.;
RL   Biochem. Biophys. Res. Commun. 148:1549-1550(1987).
RN   [7]
RP   SEQUENCE OF 1-94 FROM N.A.
```

RX   MEDLINE=91099327; PubMed=2125269;
RA   Yokoyama Y., Kawamoto T., Mitsuuchi Y., Kurosaki T., Toda K.,
RA   Ushiro H., Terashima M., Sumimoto H., Kuribayashi I., Yamamoto Y.,
RA   Maeda T., Ikeda H., Sagara Y., Shizuta Y.;
RT   "Human poly(ADP-ribose) polymerase gene. Cloning of the promoter
RT   region.";
RL   Eur. J. Biochem. 194:521-526(1990).
RN   [8]
RP   SEQUENCE OF 1-39 FROM N.A.
RX   MEDLINE=90211250; PubMed=2108670;
RA   Ogura T., Nyunoya H., Takahashi-Masutani M., Miwa M., Sugimura T.,
RA   Esumi H.;
RT   "Characterization of a putative promoter region of the human
RT   poly(ADP-ribose) polymerase gene: structural similarity to that of
RT   the DNA polymerase beta gene.";
RL   Biochem. Biophys. Res. Commun. 167:701-710(1990).
RN   [9]
RP   SEQUENCE OF 1-39 FROM N.A.
RA   Herzog H., Schneider R., Hirsch-Kauffmann M., Schnitzer D.,
RA   Schweiger M.;
RL   Submitted (JUL-1991) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=90222155; PubMed=2109322;
RA   Gradwohl G., Menissier de Murcia J., Molinete M., Simonin F.,
RA   Koken M.H.M., Hoeijmakers J.H.J., de Murcia G.M.;
RT   "The second zinc-finger domain of poly(ADP-ribose) polymerase
RT   determines specificity for single-stranded breaks in DNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:2990-2994(1990).
RN   [11]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=91072398; PubMed=2123876;
RA   Ikejima M., Noguchi S., Yamashita R., Ogura T., Sugimura T.,
RA   Gill D.M., Miwa M.;
RT   "The zinc fingers of human poly(ADP-ribose) polymerase are
RT   differentially required for the recognition of DNA breaks and nicks
RT   and the consequent enzyme activation. Other structures recognize
RT   intact DNA.";
RL   J. Biol. Chem. 265:21907-21913(1990).
RN   [12]
RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=91035460; PubMed=2121735;
RA   Simonin F., Menissier de Murcia J., Poch O., Muller S., Gradwohl G.,
RA   Molinete M., Penning C., Keith G., de Murcia G.M.;
RT   "Expression and site-directed mutagenesis of the catalytic domain of
RT   human poly(ADP-ribose)polymerase in Escherichia coli. Lysine 893 is
RT   critical for activity.";
RL   EMBO J. 11:3263-3269(1992).
RN   [13]
RP   NUCLEAR LOCALIZATION SIGNAL.
RX   MEDLINE=92371433; PubMed=1505517;
RA   Schreiber V., Molinete M., Boeuf H., de Murcia G.M.,
RA   Menissier de Murcia J.;
RT   "The human poly(ADP-ribose) polymerase nuclear localization signal is
RT   a bipartite element functionally separate from DNA binding and
RT   catalytic activity.";
RN   [14]
RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=97461532; PubMed=9315851;
RA   Rolli V., O'Farrell M., Menissier-de Murcia J., de Murcia G.M.;
RT   "Random mutagenesis of the poly(ADP-ribose) polymerase catalytic
RT   domain reveals amino acids involved in polymer branching.";
RL   Biochemistry 36:12147-12154(1997).
-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
-!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
-!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.

-!- SUBUNIT: HOMODIMER (Potential).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
-!- SIMILARITY: BELONGS TO THE PARP FAMILY.
-!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; X16674; CAA34663.1; -.
DR   EMBL; M18112; AAA60137.1; -.
DR   EMBL; J03473; AAB59447.1; -.
DR   EMBL; M17081; AAA51599.1; ALT_SEQ.
DR   EMBL; M32721; AAA60155.1; -.
DR   EMBL; M29786; AAA51663.1; -.
DR   EMBL; M29545; AAA51663.1; JOINED.
DR   EMBL; M29766; AAA51663.1; JOINED.
DR   EMBL; M29767; AAA51663.1; JOINED.
DR   EMBL; M29768; AAA51663.1; JOINED.
DR   EMBL; M29769; AAA51663.1; JOINED.
DR   EMBL; M29770; AAA51663.1; JOINED.
DR   EMBL; M29771; AAA51663.1; JOINED.
DR   EMBL; M29772; AAA51663.1; JOINED.
DR   EMBL; M29773; AAA51663.1; JOINED.
DR   EMBL; M29774; AAA51663.1; JOINED.
DR   EMBL; M29775; AAA51663.1; JOINED.
DR   EMBL; M29776; AAA51663.1; JOINED.
DR   EMBL; M29777; AAA51663.1; JOINED.
DR   EMBL; M29778; AAA51663.1; JOINED.
DR   EMBL; M29779; AAA51663.1; JOINED.
DR   EMBL; M29780; AAA51663.1; JOINED.
DR   EMBL; M29781; AAA51663.1; JOINED.
DR   EMBL; M29783; AAA51663.1; JOINED.
DR   EMBL; M29784; AAA51663.1; JOINED.
DR   EMBL; M29785; AAA51663.1; JOINED.
DR   EMBL; M29544; AAA51663.1; JOINED.
DR   EMBL; M29782; AAA51663.1; JOINED.
DR   EMBL; X56140; CAA39606.1; -.
DR   EMBL; X56141; CAA39606.1; JOINED.
DR   EMBL; M60436; AAA60000.1; -.
DR   PIR; A26901; A26901.
DR   PIR; A28498; A28498.
DR   PIR; A29725; A29725.
DR   PIR; A35635; A35635.
DR   PIR; A33321; A33321.
DR   PIR; B33321; B33321.
DR   PIR; A39976; A39976.
DR   PIR; S14010; S14010.
DR   HSSP; P26446; 1A26.
DR   Aarhus/Ghent-2DPAGE; 1620; NEPHGE.
DR   MIM; 173870; -.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.

Query Match   33.8%; Score 1012.5; DB 1; Length 1013;
Best Local Similarity   40.5%; Pred. No. 1.4e-58;
Matches 221; Conservative 111; Mismatches 185; Indels 29; Gaps 11;

Qy   29 APEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAK 88
         || | | || :: | :| .:: || | .| ||| | ||| .||::

```
Db   493  APRGKSGAALSKKSKGQV-----------KEEGINKSEKRMK-LTLKGGAAVDPD-SGL  538

Qy    89  VGKAHVYCEGNDVDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVMRWGRVGKMGQHS  148
               |||    :|   |    :  :||||||    |    :   |||||  :
Db   539  EHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSN  598

Qy   149  LVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKE  208
               :  | |||:|| :|| |   ||| |  |  :||   :|| ||
Db   599  KLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYG---QDEEAVKK-  653

Qy   209  ESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQS  268
               : |||||:::: :|:::|:: |:| | |  ::  :  | ||||||  ||:|
Db   654  --LTVNPGTKSKLPKPVQDLIKMIFDVESMKKKAMVEYEIDLQKMPLGKLSKRQIQAAYSI  711

Qy   269  LKKIEDCIRAGGHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIE  328
               : ::    ||  |||||  |||||||||||      ||||    |::|:| |||
Db   712  LSEVQQAVSQGSGSSDSQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIE  771

Qy   329  IAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMT  387
               :|:|:  |:  |:|||    ::| ||:  :|::| ||:::|| ||:  | | |
Db   772  VAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEAEIIRKYVKNTHATTHNAYDLE  831

Qy   388  LLDLFEVKDGEKEAFR--EDILHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFG  445
               ::|:||:::| |  :| ||| |  |  ||| :::|    || |||||||||  ||||||
Db   832  VIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFG  891

Qy   446  KGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGL  505
               ||||||||       |||||||||||||:||  || |  ||   ||:|||||||||||
Db   892  KGIYFADMVSKSSANYCHTSQGDPIGLILLGEVALGNMYELKHASHISK-LPRGKHSVKGL  950

Qy   506  GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFN  565
               |:           ::|: |||    : ::|||||||||||||:|::||||||||
Db   951  GKTTPDPSANISLDGVDVPLGTGISSGV---NDTSLLYNEYIVYDIAQVNLKYLLLKFN  1007

Qy   566  F-LQLW  570
          | |||||
Db  1008  FKTSLW  1013
```

RESULT 9
PPOL_BOVIN

```
ID   PPOL_BOVIN      STANDARD;       PRT;   1015 AA.
AC   P18493; Q9TS00;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90382673; PubMed=2119324;
RA   Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;
RT   "Cloning of a full-length cDNA encoding bovine thymus
RT   poly(ADP-ribose) synthetase: evolutionarily conserved segments and
RT   their potential functions.";
RL   Gene 90:249-254(1990).
RN   [2]
RP   SEQUENCE OF 647-714 AND 838-903 FROM N.A.
RX   MEDLINE=88151954; PubMed=2450019;
RA   Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,
RA   Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;
RT   "Depression in gene expression for poly(ADP-ribose) synthetase during
RT   the interferon-gamma-induced activation process of murine macrophage
RT   tumor cells.";
RL   Eur. J. Biochem. 171:571-575(1988).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N) + [ADP-D-ribosyl](N+1)-acceptor -
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D90073; BAA14114.1; -.
DR   EMBL; X06986; CAA30046.1; -.
DR   EMBL; X06987; CAA30047.1; -.
DR   PIR; JS0428; JS0428.
DR   PIR; S00328; S00328.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      1      1
FT   DNA_BIND      1    374                       BY SIMILARITY.
FT   DOMAIN      375    525                       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      386    462                       BRCT.
FT   DOMAIN      526   1015                       NAD-BINDING.
FT   ZN_FING      20     55                       PARP-TYPE.
FT   ZN_FING     127    164                       PARP-TYPE.
FT   DOMAIN      209    211                       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228                       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     408    408                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     414    414                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     446    446                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     449    449                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     457    457                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     472    472                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     489    489                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     492    492                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     514    514                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     515    515                       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     521    521                       ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE   1015 AA;  113355 MW;  0A5FE9D9F04F5B04 CRC64;

Query Match             33.7%;  Score 1009;  DB 1;  Length 1015;
```

```
Best Local Similarity   40.4%;  Pred. No. 2.3e-58;
Matches  222; Conservative  108; Mismatches  187;  Indels  32;  Gaps  12;

Qy     26 GNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPEC  85
             :  :!:!!:!:     !!        !!    :  : !!  :! : !!! ! !!!::
Db    495 GPKGKSGAAPSKKS--------KGPV------KEEGTNKSEKRMK-LTLKGGAAVDPD-  537

Qy     86 TAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMG 145
             !!! ! !   !!  ! !   !  !   :::: !  : !!!:!!!:: : :!!
Db    538 SGLEHNAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRVWIFRSWGRVGTVI 597

Qy    146 QHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEET 205
             :!  ! ! !    !  !  !!!!!: :!! ! !:!! !!::!!  !!!!!!:!!
Db    598 GSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKHPKFYPLEIDYG---QDEEAV 653

Qy    206 KKEESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAG 265
             !! -- ! :!!  !  ::! !!! :! ::!:: !: ! :! ! !!!!! ! !::!
Db    654 KK---LTVNPGTKSKLPKPVQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKILSKRQIQAA 710

Qy    266 YQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALG 325
             !!!!!!!:!   :: :  !! !!!!!!!:  !:: :!!!!!!!!: !!!:!!:!!!:!
Db    711 YSILSEVQQALSQGSSDSHILDLSNRPYTLIPHDFGMKKPPLLNNANSVQAKVEMLDNLL 770

Qy    326 DIEIAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDY 384
             !!!::!!::!  :! :!: ! :! :!!:!::::! !:! :!!  !!!!! !! :!  !
Db    771 DIEVAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVRDSERAEIIRKYVKNTHATTHNAY  830

Qy    385 TMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGY 442
             :  ::!:::!!: :: :: !!!  !!!!!! ! !!:!!: :!!! !!!!!!!!:!!!:!!!
Db    831 DLEVVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTGY 890

Qy    443 MFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHST 502
             !!!!!!!!!!:  !!!!!! !: :!!!!!! !: !!  :! !!! !  !  ! !!  !!!!
Db    891 MFGKGIYFADMVSKSANYCHTSQGDPIGLILGEAALGNMYELKHARHISK-LPKGKHSV 949

Qy    503 KGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKV 562
             !!!!!  !:  !! !  !  :  !: !       !!!!!!!:!   !!!!:!! !!! :!!
Db    950 KGLGKTPDPSASITVDGVEVPLGTGISSGV---NDTCLLYNEYIVYDIAQVHLKYLYLKL 1006

Qy    563 QFNF-LQLW 570
             !!! !  !!
Db   1007 KFNFKTSLW 1015


RESULT  10
PPOL_XENLA
ID   PPOL_XENLA      STANDARD;      PRT;      998 AA.
AC   P31669;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Saulier-Le Drean B.M.;
RL   Thesis (1992), University of Rennes, France.
RN   [2]
RP   SEQUENCE OF 742-876 FROM N.A.
RX   MEDLINE=93277538; PubMed=8503897;
RA   Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,
RA   Miwa M.;
RT   "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)
RT   polymerase from Xenopus laevis and cherry salmon using heterologous
RT   oligonucleotide consensus sequences.";
RL   Biochem. Biophys. Res. Commun. 193:119-125(1993).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,
CC       AND BRAIN. LOW IN LIVER.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z12139; CAA78126.1; -.
DR   EMBL; D13810; BAA02966.1; -.
DR   PIR; S31735; S31735.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   NON_TER       1      1
FT   DNA_BIND     <1    356
FT   DOMAIN      357    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      369    445       BRCT.
FT   DOMAIN      508    998       NAD-BINDING.
FT   ZN_FING       8     43       PARP-TYPE.
FT   ZN_FING     111    148       PARP-TYPE.
FT   DOMAIN      193    195       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      207    212       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     391    391       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     397    397       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     419    419       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     428    428       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     429    429       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     447    447       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     454    454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     467    467       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     471    471       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     477    477       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     495    495       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     496    496       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     503    503       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    746    746       Q -> E (IN REF. 2).
SQ   SEQUENCE   998 AA;  111126 MW;  F5A25E4A366BAE7 CRC64;
```

```
Query Match        32.8%;  Score 984.5;  DB 1;  Length 998;
Best Local Similarity  39.1%;  Pred. No. 9e-57;
Matches  222;  Conservative 109;  Mismatches 196;  Indels 41;  Gaps 14;

Qy    21 KRVNNGNTAPE------DSSPAKKTRRCQRQESKKMPVAG------GKANKDRTEDKQD 67
           : :|:  |:|       :  |      || ||          ::  ||:    ::|:
Db   444 KEVESGKSVQELLSQFGISSWGAEIKQEAVQPTEKQPSSGPVAGKSSGKVKEEKGSNKSE 503

Qy    68 ESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDD 127
          ::|:|:|:  |:||:  ||    ||     |||     |  :::  |:::| :|:|:||
Db   504 KKMK-LTVKGGAAIDPDSELE-DSCHVLETGGKIFSATLGLVDITRGTNSYKLQLIEHD 561

Qy   128 AQRNFSVWMRWGRVGK-MGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPG 186
          :  |:   |||||  | :|   |   :|  |   |  |||  | | |    |  |   ||
Db   562 RDSRYWVFRSWGRVGVTVIGSKKLEEMSSK-EDAIEHFLNLYQDKTGNAWHS-PNFTKYPK 619

Qy   187 KYDMLQMDYATNTQDEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMMEMK 246
          |: :|||  |: |: | : |  | |:|   || : | |||  |:||||   || |:|:|
Db   620 KFYPLEIDYG---QEEDVVKK---LSVGAGTKSKLAKPVQELIKLIFDVESMKKAMVEFE 673

Qy   247 YNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEYTRIPHDFGLRTPP 306
          :  :||||||  |:  |||  ||||:    |||||||||:|  ||:||||||:|||:||
Db   674 IDLQKMPLGKLSKRQIQSAYSILSQVQQAVESLSEARLLDLSNQFYTLIPHDFGMKKPP 733

Qy   307 LIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPE-HPLDQHYRNLHCALRPLDHESYE 365
          |: :|: : |:||: |:|||  | ||| |:|| |:  ||  | |:|:|   | |: |::
Db   734 LLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDGEKDPIDVKYEKIKTDIKVVAKDSEE 793

Qy   366 FKVISQLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWV 423
          :|: : |      :|    || || :|:|| |  |     |:|||:| ||||||: |:|
Db   794 SRIICDYVKNTHADTHNAYDLEVLEIFKIDREGEYQRYKPFKQLHNRQLLWHGSRTTNFA 853

Qy   424 GILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCN 483
          ||||:||||||||| |||||||||||||||:|||||||  |||||||||||||||||||
Db   854 GILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANYCHAMPGSPIGLILLGEVALGNMH 913

Qy   484 ELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPA-----SDTGILNPDG 538
          |||:||  ||||:||  | |||| ||        |||  ||||||     || ||| |
Db   914 E-LKAASQITKLPKGKHSVKGLGRTRAPDPSATVQLDGVDVPLGKGTSANISDTSLL---- 968

Qy   539 YTLNYNEYIVYNPNQVRMRYLLKVQFNF 566
          |||| |||| ||  || ||||||| ||
Db   969 ----YNEYIVYDIAQVNLKYLLKLKFNY 992
```

```
PPOL_DROME
ID   PPOL_DROME     STANDARD;     PRT;     994 AA.
AC   P35875; Q9W5Q5; Q9W5S1;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   PARP OR CG17696/CG17718.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93234521; PubMed=8475096;
RA   Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,
RA   Sugimura T., Miwa M.;
RT   "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:
RT   leucine zipper in the auto-modification domain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND
RP   TISSUE SPECIFICITY.
RC   STRAIN=CANTON-S;
RX   MEDLINE=98234380; PubMed=9565614;
```

RA   Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;
RT   "Genomic organization of Drosophila poly(ADP-ribose) polymerase and
RT   distribution of its mRNA during development.";
RL   J. Biol. Chem. 273:11881-11886(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).

```
-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC     PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC     ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC     CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC     TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC     EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC  -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC     nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC  -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC  -!- SUBCELLULAR LOCATION: Nuclear.
CC  -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS;
CC     SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC  -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES
CC     OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM
CC     IN LATER EMBRYOS.
CC  -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN
CC     EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.
CC  -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC     AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC     FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC     THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC     AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC  -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC  -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   ---------------------------------------------------------------------------
```

```
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; D13806; BAA02964.1; -.
DR   EMBL; AF051548; AAC24518.1; -.
DR   EMBL; AF051544; AAC24518.1; JOINED.
DR   EMBL; AF051545; AAC24518.1; JOINED.
DR   EMBL; AF051546; AAC24518.1; JOINED.
DR   EMBL; AF051547; AAC24518.1; JOINED.
DR   EMBL; AE002935; AAF45400.1; -.
DR   EMBL; AE002666; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE002892; AAF45445.2; ALT_SEQ.
DR   PIR; A47474; A47474.
DR   HSSP; P26446; 1A26.
DR   FlyBase; FBgn0010247; Parp.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.
FT   DNA_BIND      1    367
FT   DOMAIN      368    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      380    454       BRCT.
FT   DOMAIN      508    994       NAD-BINDING.
FT   ZN_FING      19     54       PARP-TYPE.
FT   ZN_FING     123    161       PARP-TYPE.
FT   DOMAIN      208    210       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   VARSPLIC    376    564       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   994 AA;  113791 MW;  ACA85A270DD29E08 CRC64;
```

```
Query Match             32.7%;  Score 979.5;  DB 1;  Length 994;
Best Local Similarity   38.3%;  Pred. No. 1.9e-56;
Matches  215;  Conservative 110;  Mismatches 199;  Indels 37;  Gaps 11;

Qy     12  GRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQE--SKKMPVAGGKANKDRTEDKQDES   69
               | ::  :::||| : :   :|::| |:::   |  : |||:
Db    463  GAIKYIN-STCICSWGTDPKSRIPKETTKSLNSNSIYTKSMPVS----------------  505

Qy     70  VKALLLKGKAPVDPPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQ  129
               ||::  :   ::|  |:|:| ||:|||| | :|| |:: :| |
Db    506  -RTFKVKDGLAVDPD-SGLEDIAHVYVDSNNKYSVVLGLTDIQRNKNSYYKVQLLKADKK  563

Qy    130  RNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYD  189
               || |:: ::: ||:|: |  : |: |||: :|| |:|:: ||| |
Db    564  EKYWIFRSWGRIGTNIGNSKLEEFPTSESAKRNFKEIYADKTGNEYEQRDNFVKRTGRMY  623

Qy    190  MLQMDYATNTQDEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEEMMMKYNT  249
                ||:  :|| ||| :  :||:: : :|:|:|:|: |:: ||||
Db    624  PIEIQY---DDDQKLVKHESHFF----TSKLEISVQNLIKLIFDIDSMNKTLMEFHIDM  675

Qy    250  KKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIR  309
               | || |:: : ||:: |:|:|||:|:|:: :|::| || |
Db    676  DKMPLGKLSAHQIQSAYRVVKEIYNVLECGSNTAKLIDATNRFYTLIPHFGVQLPTLIE  735

Qy    310  TQKELSEKIQLLEALGDIEIAIKLVKTE-LQSPEHPLDQHYRNLHCALRPLDHESYEFKV  368
               |::|:|:|| |:|:|  :| : ::   |:|:|:  ||||| |:::| : ||::
Db    736  THQQIEDLRQMLDSLAEIEVAYSIIKSEDVSDACNPLDNHYAQIKTQLVALDKNSEEFSI  795
```

```
Qy    369  ISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGIL  426
               :||::|||: ||: ::::::| ::::| ||||::|  |  |::|||||||:::|||
Db    796  LSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARRFKPFKLHNRKLLWHGSRLTNFVGIL  855

Qy    427  SHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLSEVALGQCNELL  486
               ||||||||||| ||||||||||||| | ::|||| || ||||| || ::|||  |:||
Db    856  SHGLRIAPPEAPTGYMFGKGIYFADMVSKSANYCCTSQQNSTGLMLLSEVALGDMMECT  915

Qy    487  EANPKAEGLLQGKHSTKGLGKMAPS-SAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNE  545
               || ::| | ||::||:|:|||:||: ||::: |||::||| ::: : ||||||||||
Db    916  SAK-YINKLSNNKHSCFGRGRTMPDPFKSYIRSDGVEIPYGETITDEHLKS---SLLYNE  971

Qy    546  YIVYNPNQVRMRYLLKVQFNF  566
               ||::|  :||:|  ::|
Db    972  YIVYDVAQVNIQYLFRMEFKY  992
```

```
RESULT 12
PPOL_SARPE
ID   PPOL_SARPE     STANDARD;      PRT;    996 AA.
AC   Q11208;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
OS   Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7386;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=94170813; PubMed=8125121;
RA   Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,
RA   de Prati A.C., Kurata S., Natori S., Suginura T., Esumi H.;
RT   "Cloning and functional expression of poly(ADP-ribose) polymerase
RT   cDNA from Sarcophaga peregrina.";
RL   Eur. J. Biochem. 220:607-614(1994).
-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics  and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; D16482; BAA03943.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
```

```
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; FALSE_NEG.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; zinc-finger; Zinc.
FT   DNA_BIND      1    369       BY SIMILARITY.
FT   DOMAIN      370    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      382    456       BRCT.
FT   DOMAIN      508    996       NAD-BINDING.
FT   ZN_FING      19     54       BY SIMILARITY.
FT   ZN_FING     126    164       BY SIMILARITY.
FT   DOMAIN      211    214       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      232    235       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
SQ   SEQUENCE    996 AA;  113018 MW;  690DDD36E7487298 CRC64;

Query Match             32.2%;  Score 964;  DB 1;  Length 996;
Best Local Similarity   39.9%;  Pred. No. 2e-55;
Matches  216;  Conservative 116;  Mismatches 177;  Indels  32;  Gaps 13;

Qy    33  SSPAKKTRRCQRQESKKMPVAGGKANKDRTEDKDQDESVKALLLKGKAPVDPECTAKVGKA  92
Db   480  TDPATRITQEESKSSSKSKSIYTKSVPKSMT-------LKIKDGLAVDPD-SGLEDVA   528

Qy    93  HVYCEGN-DVYDVMLNQTNLQFNNNKYLIQLLEDDAQRNFSVWMRWGRVG-KMGQHSLV  150
Db   529  HVYVSRNKEKYNVLGITDIQKNKNSFYKLQLLESDMKNRFWFRSWGRIGTTIGGNKLD  588

Qy   151  ACSGNLNAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEES  210
Db   589  NFS-NLVDAIVQFKELYLEKSGNHFENRENFVKVAGRMYPIDIDYA------EDSKIDLS  641

Qy   211  LKSPLKPESQLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLK  270
Db   642  AEHDIK--SKLPLSVQDIIKLMFPVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLT  699

Qy   271  KIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLLRTQKELSEKIQLLEALGDIEIA  330
Db   700  EIYELIQGGGTNAKFIDATNRFYTLIPHNFGTQSPPLLDTTEQVEQLRQMLDSLIEIECA  759

Qy   331  IKLVKTELQSPE-HPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLL  389
Db   760  YSLLQTEDSKADINPIDKHYEQLKTKLEPLDKNSEEYILLQKVVKNTHAETHKLYDLEVV  819

Qy   390  DLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWGILSHGLRIAPPEAPITGYMFGKG  447
Db   820  DIFKVARQGEARRYKPFFKLHNRRLLWHGSRLNFAGILSHGLKIAPPEAPVTGYMFGKG  879

Qy   448  IYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGK  507
Db   880  IYFADMVSKSANYCCTSHHNSTGLMLLSEVALGDMMECTAAKYVTK-LPNDKHSCFGRGR  938

Qy   508  MAPS-SAHFVTLNGSTVPLG-PASDTGILNPDGYTLNVEYIVYNPNQVRMRYLLKVQFN  565
Db   939  TMPNPSESIIREDGVEIPLGKPITNDSLKS----SLLYNEFIIYDIAQVNIQYMLRMNFK  994

Qy   566  F 566
Db   995  Y 995

RESULT 13
PPO3_HUMAN
ID   PPO3_HUMAN      STANDARD;       PRT;      533 AA.
AC   Q9Y6F1; Q9UG81;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+) ADP-
```

```
DE   ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (pADPRT-3)
DE   (hPARP-3).
GN   ADPRTL3 OR PARP3 OR ADPRT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=9263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [2]
RP   SEQUENCE OF 75-533 FROM N.A.
RC   TISSUE=Kidney;
RA   Ansorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels are in
CC       the kidney, skeletal muscle, liver, heart and spleen; also
CC       detected in pancreas, lung, placenta, brain, leukocytes, colon,
CC       small intestine, ovary, testis, prostate and thymus.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF083068; AAD29855.1; -.
DR   EMBL; AL050034; CAB43246.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation.
FT   DOMAIN       14     20       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   CONFLICT     80     80       K -> N (IN REF. 2).
FT   CONFLICT    171    171       A -> G (IN REF. 2).
FT   CONFLICT    411    411       K -> E (IN REF. 2).
SQ   SEQUENCE    533 AA;  60117 MW;  7C0AB89E64D1B9FD CRC64;

Query Match             24.0%;  Score 720;  DB 1;  Length 533;
Best Local Similarity   34.8%;  Pred. No. 7.3e-40;
Matches  193;  Conservative  98;  Mismatches  188;  Indels  76;  Gaps  20;

Qy    46  ESKKMPVAGGKANKDRTED---KQDESVKALLLKGK-APVDPECTAKVGK-AHYCEGND  100
Db    15  EKKKGRQAG------REEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVY----E   64

Qy   101  VYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAK  160
Db    65  DYNCTLNQTNIENNNKKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDAK  122

Qy   161  EIFQKKFLDKTKNNWEDREKFPEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ  220
Db   123  KDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVKVDRAPVRTVTK  178

Qy   221  ------LDLRVQELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIED  274
Db   179  RVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEE  238
```

```
Qy   162 IFQKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNT-QDEETKKEE------------ 209
             :: :: .:: :  ::  ::. ::  :  :: : ||: ::: ||| .||
Db    69 FFESKFYEKTHLHWEERDD-EPVPNKYAVVEL--ATNARQTEKEVKKEEPEPKVDEKN   125

Qy   210 -----------SLKSPLKPESQLDLRVQELIKLICNVQAMEMMMEMKYNTK-KAPL    254
                        :: :. :. ||: :  : :  : :: :  :  :  . :|: |:::
Db   126 TRGRKKRGIVKEKKEIKKEEPVEEVNEKLKELMKCICDEDVHLGLLKQLKFNEAFGRPI  185

Qy   255 GKLTVAQIKAGYQSLKIEDCI-----RAGQHGR-------------ALMEACNEFYTR   295
             .::.||| :||  .|| :.|          :|:              ||||||||||
Db   186 DCLSLAQLTTGYEILSKIEESIGGKSARRSTRGRPRVADRVLAVKSDGPSLHDINKYYSL 245

Qy   296 IPHDFGLRTPPLIRTQKELSEKIQLLEAL-GDIEIAIKLVKTELQSPEHPLDQH-YRNLH  353
             ||| :| | |::::: |: :::||:  :: .::. : : :.:::: :  ||::  ||||
Db   246 IPHSFGFCVPPKIDSHAKIQAERELLDALKGSIEASLELKDLKKTASSKDIYQRLYERLP  305

Qy   354 CALRPLDHESYEFKVISQYL--QSTHAPTHSDYTMTLLDLFEVEKDGE----------   399
             | | ::   .:  :  ::      ::    . :  : :: :   :.
Db   306 CHLEPVSEE------IAGKIGDCLAMRGPTHC-YKLSLIDAFELKDPNEIPTEAPVEVQE  358

Qy   400 --KEAFREDLH----------NRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKG  447
               : ::  :           :: ||||||||||||||||||| :|||||||||||||
Db   359 VPKKRGRKSTKTAAPTVPPPTKRLLWHGTRVNVFSILMNGLQF--PVGDRCGLMFGNG    416

Qy   448 IYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQ--GKHSTKGL  505
             :::|. | |||||| . | .:  .::| .|||.||| | |:|    :      ||.| |:
Db   417 VFANVPFKSANYCCPEASKRV-FMLLCEVETANPLVLYESEIDADEKMEKAKKTSVVAA   475

Qy   506 GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYT-LNYNEYIVYNPNQVRMRYLLKVQF  564
             ||  ||:||||| |||.|| |||||:|||| || | |||||:|::||||||||||||:||
Db   476 GKHTPRDT--VEINGI------PAFKSNLETIEEETRLLYDEYVMFNKEHFKIKYVVEVKV 528

Qy   565 NFL 567
             :
Db   529 DRL 531
```

```
Qy   275 CIRA-GQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEI--AI  331
             ::   .:.|:|||||||||| ||| :| | |::::: |: :::||:  :: .::.  ::
Db   239 ALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQAL  298

Qy   332 KLVKTELQSPE--HPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTL   388
             :.::.|:.:.    ||: | | | | | ::   .:  :  ::  .  ::   :  :  :|
Db   299 QAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS--NHRCPTL   355

Qy   389 LDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRITAPPEAPITGYMFGK 446
             :|:  :|  :. .::     :: ||||||| . :.|||:| :.|| ::  .:. ||.|
Db   356 QHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGK    411

Qy   447 GIYFADMSSKSANYCFASR-LKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKG   504
             |||||| |.||| |::| | | .:  ::: .|||| ::: :| .: :   ::
Db   412 GIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL----------    459

Qy   505 LGKMAPSSAHFVTLNGSTVPLGPASDT--------GILNPDGY-----------TLNYNEYI 547
             :||| | |.   .. |:||. | :.:|          |::|.            ||||||
Db   460 --KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYL    516

Qy   548 VYNPNQVRMRYLLKV 562
             :|:.: | |||||:|
Db   517 IYQESQCRLRYLLEV 531
```

RESULT 14
YQN4_CAEEL
ID   YQN4_CAEEL   STANDARD;   PRT;   538 AA.
AC   Q09525;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.
GN   E02H1.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Smith A.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-
CC   RIBOSYLTRANSFERASE (EC 2.4.2.30).
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; Z47075; CAA87379.1; -.
DR   HSSP; P26446; 1A26.
DR   WormPep; E02H1.4; CE01539.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   538 AA; 61268 MW; 3144E25465FC7341 CRC64;

Query Match   17.5%; Score 526; DB 1; Length 538;
  Best Local Similarity   29.1%;
  Matches 158; Conservative 103; Mismatches 182; Indels 100; Gaps 22;

```
Qy   102 YDVMLNQTNLQFNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKE  161
             | |:| :||:||||.:| :: ::|:   :|: ::|||:||: ||  ::| :|:|:||:
Db    12 YKVHLCKTNIAQNNNKFYDMELLDEGG--DFIVKLINGRIGYRGVTQLKFDD-DLDRAKK  68
```

RESULT 15
PPOL_ONCMA
ID   PPOL_ONCMA   STANDARD;   PRT;   135 AA.
AC   Q08824;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).
GN   ADPRT.
OS   Oncorhynchus masou (Cherry salmon) (Masu salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8020;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93277538; PubMed=8503897;
RA   Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,
RA   Miwa M.;
RT   "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)
RT   polymerase from Xenopus laevis and cherry salmon using heterologous
RT   oligonucleotide consensus sequences.";
RL   Biochem. Biophys. Res. Commun. 193:119-125(1993).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY[ADP-RIBOSYL]ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,
CC   AND BRAIN. LOW IN LIVER.

```
CC  -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC      AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC      FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC      THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC      AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC  -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC  ---------------------------------------------------------------------
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute.  There are no restrictions on its
CC      use by non-profit institutions as long  as its content is in no way
CC      modified and this statement is not removed.  Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC  ---------------------------------------------------------------------
DR  EMBL; D13809; BAA02965.1; -.
DR  PIR; PN0494; PN0494.
DR  HSSP; P26446; 1A26.
DR  InterPro; IPR001290; PARP.
DR  InterPro; IPR001510; Znf-PARP.
DR  Pfam; PF00644; PARP; 1.
DR  PROSITE; PS00347; PARP_ZN_FINGER_1; PARTIAL.
DR  PROSITE; PS50064; PARP_ZN_FINGER_2; PARTIAL.
KW  Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW  ADP-ribosylation; Zinc-finger; Zinc.
FT  NON_TER       1       1
FT  DOMAIN       <1     >135        NAD-BINDING.
FT  ACT_SITE     135     135        BY SIMILARITY.
FT  NON_TER      135     135
SQ  SEQUENCE  135 AA;  15411 MW;  A60E8E98890F42DC CRC64;

Query Match              11.1%;  Score 331.5;  DB 1;  Length 135;
Best Local Similarity    46.6%;  Pred. No. 2.7e-15;
Matches   62;  Conservative   34;  Mismatches   34;  Indels    3;  Gaps    2;

Qy    317 KIQLLEALGDIEIAIKLVKTELQ-SPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQS 375
          |:::|:|||:|| |:|  : :|:||:|:|  :  : :|:|:  | :| |:::
Db      3 KVEMLDNLLDIEVAYSLLKGGAEDNKKDPIDINYEKLKTKIEVVDKTTKEAEIILQYVKN 62

Qy    376 THAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIA 433
          ||| ||: ||:: | : :::| : ::|   | || ||||| :|:| ||| |||||
Db     63 THAATHNTYTLVVEEIFKIVREGEYQKYRPFQDLPNRQLLWHGSRATNYAGILSQGLRIA 122

Qy    434 PPEAPITGYMFGK 446
          |||||:||||||
Db    123 PPEAPVTGYMFGK 135
```

Search completed: August 29, 2002, 08:02:11
Job time: 288 sec

```
              GenCore version 4.5
        Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     August 29, 2002, 07:55:28 ;  Search time 117.38 Seconds
                                          (without alignments)
                                          840.067 Million cell updates/sec

Title:        US-09-701-586B-2
Perfect score: 2998
Sequence:     1 MAARRRRSTGGGRARALNES.........PNQVRMRYLLKVQFNFLQLW 570

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:     562222 seqs, 17994929 residues

Total number of hits satisfying chosen parameters:     562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_19:*
         1:  sp_archea:*
         2:  sp_bacteria:*
         3:  sp_fungi:*
         4:  sp_human:*
         5:  sp_invertebrate:*
         6:  sp_mammal:*
         7:  sp_mhc:*
         8:  sp_organelle:*
         9:  sp_phage:*
        10:  sp_plant:*
        11:  sp_rodent:*
        12:  sp_virus:*
        13:  sp_vertebrate:*
        14:  sp_unclassified:*
        15:  sp_rvirus:*
        16:  sp_bacteriap:*
        17:  sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1176	39.2	635	10	081294	081294 arabidopsis
2	1149.5	38.3	653	10	050017	050017 zea mays (m
3	1022	34.1	607	13	09PS82	09ps82 gallus gall
4	1018	34.0	1014	11	09Z1K2	921k2 mus musculu
5	988	33.0	612	5	Q9TX06	09tx06 dictyosteli
6	981.5	32.7	607	13	Q9PS81	09ps81 xenopus. na
7	972.5	32.4	593	5	Q9TX05	09tx05 drosophila
8	966	32.2	983	10	Q9ZP54	09zp54 arabidopsis
9	966	32.2	1009	10	Q9SJW4	09slw4 arabidopsis
10	961.5	32.1	969	10	024570	024570 zea mays (m
11	961.5	32.1	980	10	Q9ZSV1	09zsv1 zea mays (m
12	777.5	25.9	945	5	Q9N4H4	09n4h4 caenorhabdi
13	725	24.2	533	4	Q96CG2	96cg2 homo sapien
14	696.5	23.2	528	11	Q91YR6	091yr6 mus musculu
15	685.5	22.9	727	5	Q9XUA5	09xua5 caenorhabdi
16	408.5	13.6	815	10	Q9FK91	09fk91 arabidopsis
17	396.5	13.2	815	10	Q9SWB4	Q9swb4 glycine max
18	353.5	11.8	2276	5	Q9TXQ1	Q9txq1 caenorhabdi
19	141.5	4.7	181	12	055721	055721 chilo iride
20		4.5	135	6	Q95JR0	095jr0 macaca fasc
21		4.5	135		095271	095271 homo sapien
22	134.5	4.5	1327		Q9X0R4	Q9x0r4 thermotoga
23	134.5	4.5	1170	16	Q95XK2	095xk2 caenorhabdi
24	133.5	4.5	1933	5	Q9AHK9	Q9ahk9 borrelia bu
25	128.5	4.3	1065	5	027421	027421 drosophila
26		4.3	1553		Q9XZ37	09xz37 drosophila
27	128	4.3	1181	5	Q9VBP3	09vbp3 drosophila
28	127.5	4.3	1181	5	Q9LAX2	09lax2 streptococc
29	125.5	4.2	479	2	Q9LAX5	09lax5 streptococc
30		4.2	481	2	049546	049546 mycoplasma
31	125	4.2	962		049524	049524 mycoplasma
32	125	4.2	1051		Q9AHK6	09ahk6 borrelia bu
33	123	4.1	1365	2	Q9AHL1	09ahl1 borrelia bu
34	122.5	4.1	1173		Q9BK46	09bk46 plasmodium
35	122.5	4.1	957	5	Q9AHK8	09ahk8 borrelia bu
36	122	4.1	3130		Q95JY2	95jy2 macaca fasc
37	121	4.0	1065	6	009471	009471 homo sapien
38	121	4.0	533		Q9H2K2	09h2k2 homo sapien
39	121	4.0	753	4	Q9NTC1	09ntc1 homo sapien
40	121	4.0	1166	4	Q9HAS4	09has4 homo sapien
41	121	4.0	1206	4	Q96Q89	96q89 homo sapien
42	121	4.0	1265		000467	000467 homo sapien
43	120.5	4.0	1820	4	Q20749	020749 caenorhabdi
44	120.5	4.0	469	5	Q910C5	0910c5 gallus gall
45	120.5	4.0	1270	13		

ALIGNMENTS

```
RESULT   1
081294
ID   081294        PRELIMINARY;      PRT;   635 AA.
AC   081294;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE).
GN   T14P8.19 OR AT4G02390.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   WASHU;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Kalicki J., Elliott G., Cloud J.;
RT   "The sequence of A. thaliana T14P8.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Waterston R.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
```

```
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babyichuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants possess two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ222588; CAA10888.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
SQ   SEQUENCE   653 AA;  72995 MW;  5FD01923C4ABCD1D CRC64;

Query Match             38.3%;       Score 1149.5;  DB 10;  Length 653;
Best Local Similarity   44.0%;       Pred No. 4.2e-72;
Matches  255;  Conservative   89;  Mismatches  191;  Indels  45;  Gaps  13;

Qy      1 MAARRRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCRQESKKMPVAGGKANKD 60
            :  :    ::     :||| ::       ::         :  : || ||
Db    102 LAKARGVAANGGKKDVIQRLLSATAGPAAVADGGPL-----------GAKEVIKGG---DE 148

Qy     61 RTEDKQDESVKALLLKGKAPVDPECTAKVG-KAHVYCEGNDVYDVMLNQTNLQFQNNNKYY 119
             |::| ::||  | || ||  :                ||:||  ||||||   ::|| :
Db    149 EVEVKKEKMVTA-TKKGAAVLDQHIPDHIKVNHVLQVGDEIYDATLNQTNVGDNNNKFY 207

Qy    120 LIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE 179
          :|| :|    ::| :| ||||||| ||  |       : || : :| |::| :| | :|
Db    208 IIQVLESDAGGSFMVYNRWGRVGVRGQDKLHGPSPTRDQAIYEFEGKFHNKTNNHWSDRK 267

Qy    180 KFEKVPGKYDMLQMDYATNTQDEEETKKE--ESLKSPLKPESQLDLRVQELIKLICNVQ 236
            : :     :  |||| |   ::: ::|      ::|  :  :  :  : :|:|||| :
Db    268 NFKCYAKKYTWLEMDYG------ETEKEIEKGSITDQIK-ETKLETRIAQFISLICNIS 319

Qy    237 AMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRI 296
           | ::: | :||   ||||||  :   |||  |||| | ||| || |  : ::  ||||  |
Db    320 MMKQRMVEIGYNAEKLPLGKLRKATILKGYHVLKRISDVISKADR-RHLEQLTGEFYTVI 378

Qy    297 PHDFGLRTPP--LIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHC 354
          |||||  ::    | : |:|  ::: ||||| |||| |||: ::| : | |  : | :|||
Db    379 PHDFGFRKMREFIIDTPQKLKAKLEMVEALGEIEIATKLLEDDSSDQDDPLYARYKQLHC 438

Qy    355 ALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE--DLHNRML 412
             |:|| :|  :|  :||| |||  ||| ||:: :: | : :: || || :  : | |||
Db    439 DFTPLEADSDEYSMIKSYLRNTHGKTHSGYTVDIVQIFKVSRHGETERFQKFASTRNRML 498

Qy    413 LWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLL 472
          |||||: ||| |||| |||||||||||||||||||:||||| ||||||| |:   :| :|
Db    499 LWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGKGVYFADMFSKSANYCYASEACRSGVL 558

Qy    473 LLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHF-VTLNGSTVPLG-----526
          || ||||| ||||| |   |   | ||: :||:|: || ::  :     | |:|||
Db    559 LLCEVALGDMNELLNADYDANNLPKGKLRSKGVGQTAPNMVESKVADDGVVVPLGGEPKQE 618

Qy    527 PASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNF 566
          |:       :   :||||||||||  |:|||| | |::|||
Db    619 PSKRGGLL-------YNEYIVYNVDQIRMRYVLHVNFNF 650

RESULT 3
Q9PS82
ID   Q9PS82   PRELIMINARY;   PRT;   607 AA.
AC   Q9PS82;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Gallus gallus (Chicken).
```

```
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF069298; AAC19283.1; -.
DR   EMBL; AL161494; CAB80732.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
KW   Transferase.
SQ   SEQUENCE   635 AA;  72017 MW;  E3F1CBE4D367A377 CRC64;

Query Match             39.2%;
Best Local Similarity   43.9%;
Matches  254;  Conservative   98;  Mismatches  187;  Indels  40;  Gaps  11;

Qy      2 AARRRRSTGGGR-----ARALNESKRVNNGNTAPEDSSPAKKTRRCRQESKKMPVAGGKA 57
            : :   || :: :||||:: 
Db     81 AIKRGLDTTGTKKDLLERLCNDANNVSN--APVKSS-------------------NDEA 118

Qy     58 NKDRT--EDKQDESVKALLLKGKAPVDPECTAKV-GKAHVYCEGNDVYDVMLNQTNLQF 113
           :|||   |:|::| : || || ||  :             : ||:||  ||||||
Db    119 EDDNNGFEEEKEEKIVTATKKGAAVLDQWIPDEIKSQYHVLQRGDDVYDAILNQTNVRD 178

Qy    114 NNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKN 173
          |||| : :| :|    : | :| ||||||| ||  :       : || : || |::|:||
Db    179 NNNKFFVLQVLESDSKKTYMVYTRWGRVGVKGQSKLDGPYDSWDRAIEIFTNKFNDKTKN 238

Qy    174 NWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPE-SQLDLRVQELIKLI 232
           | :|:  : :     :  |||| |   ::: ::|:   ::|  :  :  :  :
Db    239 YWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNNDIPSSSSEVVKFISLI 298

Qy    233 CNYQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEF 292
          || : | ::| :||   ||||||  :   |||  |||| | ||| || |  : ::  |||
Db    299 CNVSMMAQHMMEIGYNANKLPLGKISKSTISKGYEVLKRISEVIDRYDRTR-LEELSGEF 357

Qy    293 YTRIPHDFGLR--TPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYR 350
          || ||||||  :    | : |:|  ::: ||||| |||| |||:
Db    358 YTVIPHDFGFRKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLLSVDPGLQDDPLYYHYQ 417

Qy    351 NLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFRE--DLH 408
          ||||  |:|| :|  :|  :||| |||  ||| ||:: :: | : :: || || :  : |
Db    418 QLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVEIAQLFRASRAVEADRFQQFSSSK 477

Qy    409 NRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKN 468
          |||||||||: ||| |||| |||||||||||||||||||:||||| |||||||
Db    478 NRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKGVYFADMFSKSANYCYANTGAN 537

Qy    469 TGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTL-NGSTVPLGP 527
           | :|| ||||| ||||| |   |   | ||: :||:|: || ::  :     | |:||
Db    538 DGVILLCEVALGDMNELLYSDYNADNLPPGKLSTKGVGKTAPNPSEAQTLEDGVVVPLGK 597

Qy    528 ASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNF 566
          |:       :   :||||||||||  |:|||| | |::|||
Db    598 PVERSCSKG---MLLYNEYIVYNVEQIKMRYVVQVKFNY 633

RESULT 2
O50017
ID   O50017   PRELIMINARY;   PRT;   653 AA.
AC   O50017;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   607 AA;  68033 MW;  75F6E1D30D8F402 CRC64;

Query Match             34.1%;  Score 1022;  DB 13;  Length 607;
Best Local Similarity   40.6%;  Pred. No. 3.1e-63;
Matches  217;  Conservative  110;  Mismatches  187;  Indels  20;  Gaps  11;

Qy   41 RCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVCEGND 100
Db   90 KCSKPANMK---SAGKVKEEQPSKSEKKMK-LTVKGGAAVDPD-SGLEDSAHVFEKGGK 144

Qy  101 VYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVMRWGRVGKMGQHSLVACSGNLNKAK 160
Db  145 IFSATLGLVDIVKGTNSYYKLQLLEDDRESRYWFRSWGRVGTVIGSNKLEQMPSKEDAV 204

Qy  161 EIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ 220
Db  205 EHFLNLYEEKTGNSWHSK-NFTKYPKFKFYPLEIDYG---QDEEAVRK---LTVSAGTKSK 257

Qy  221 LDLRVQELIKLICNVQAMEEMMEMKYNTKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQ 280
Db  258 LAKPIQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILNEVQQAVSDGG 317

Qy  281 HGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLEAL... 339
Db  318 SESQILDLSNRFYTLIPHDFGMKKPPLLSNLEYIQAKVQMLDNLLDIEVAYSLLRGGNED 377

Qy  340 SPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE 399
Db  378 GDKDPIDINYEKLRTDIKVVDKDSEEAKIIKQYVKNTHAATHNAYDLKVVEIFRIEREGE 437

Qy  400 KEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKS 457
Db  438 SQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKS 497

Qy  458 ANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVT 517
Db  498 ANYCHTSQADPIGLILLGEVALGNMYELKNASHITK-LPKGKHSVKGLGKTTPDPTATTT 556

Qy  518 LNGSTVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNF-LQLW 570
Db  557 LDGVEVPLGNGISTGI---NDTCLLYNEYIVYDVAQVNLKYLLKLKFNYKTSLW 607

RESULT 4
Q921K2
ID   Q921K2;   PRELIMINARY;   PRT;   1014 AA.
AC   Q921K2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE)
```

```
DE   POLYMERASE).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC012041; AAH12041.1; -.
KW   Transferase.
SQ   SEQUENCE   1014 AA;  112721 MW;  5DCE68E4CB3F46EB CRC64;

Query Match             34.0%;  Score 1018;  DB 11;  Length 1014;
Best Local Similarity   41.0%;  Pred. No. 1.2e-62;
Matches  227;  Conservative  106;  Mismatches  176;  Indels  44;  Gaps  15;

Qy   29 APEDSS--PAKKTRRCQRQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECT 86
Db  495 APKGKSAAPSKKSKGAVKEEG----------VNKSEKRMK-LTLKGGAAVDPD-S 537

Qy   87 AKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYLYIQLLEDDAQRNFSVWMRWGRVGKMGQ 146
Db  538 GLEHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKRESRYWIFRSWGRVGTVIG 597

Qy  147 HSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETK 206
Db  598 SNKLEQMPSKEDAVEHFMKLYEEKTGNAWHSK--NFTKYPKKFYPLEIDYG---QDEEAVK 653

Qy  207 KEESLKSPLKP--ESQDLRVQELIKLICNVQAMEEMMEMKYNTKAPLGKLTVAQIKA 264
Db  654 ----KLTVKPGTKSKLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQIQA 708

Qy  265 GYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEAL 324
Db  709 AYSILSEVQQAVSQGSSESQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNL 768

Qy  325 GDIEIAIKLVK-TELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSD 383
Db  769 LDIEVAYSLLRGGSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTHNA 828

Qy  384 YTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITG 441
Db  829 YDLEVIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAPPEAPVTG 888

Qy  442 YMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHS 501
Db  889 YMFGKGIYFADMVSKSANYCHTSQGDPIGLILLGEVALGNMYELKHASHISK-LPKGKHS 947

Qy  502 TKGLGKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYT---LNYNEYIVYNPNQVRMRY 558
Db  948 VKGLGKTTPDPSASITLEGVEVPLG----TGI--PSGVNDTCLLYNEYIVYDIAQVNLKY 1001

Qy  559 LLKVQFNF-LQLW 570
Db 1002 LLKLKFNFKTSLW 1014

RESULT 5
Q9TX06
ID   Q9TX06;   PRELIMINARY;   PRT;   612 AA.
AC   Q9TX06;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
```

Left column:

```
RA    Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA    Kofler B., Schweiger M., Wagner E.F.;
RT    "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT    ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT    inactivation of the ADPRT gene in the mouse.";
RL    Biochimie 77:444-449(1995).
DR    HSSP; P26446; 1A26.
DR    InterPro; IPR001290; PARP.
DR    InterPro; IPR004102; PARP_reg.
DR    Pfam; PF00644; PARP; 1.
DR    Pfam; PF02877; PARP_reg; 1.
SQ    SEQUENCE  612 AA;  69241 MW;  CB340F7A88FF2364 CRC64;

Query Match            33.0%;   Score 988;  DB 5;  Length 612;
Best Local Similarity  39.6%;   Pred. No. 7.4e-61;
Matches  230;  Conservative 100;  Mismatches 207;  Indels 44;  Gaps 16;

Qy    1  MAARRRSTGGGRARALNESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKANKD  60
                   :||     ||:  ::  : :  ||  ||||         :
Db   57  VAKPTPRRSSKKAPAKSEEKGEDKNKLTT----RPTKKTKTNTSEAETKTAASDLDDSSS  111

Qy   61  RTEDKQDE-SVKALLLKGAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYY  119
             ||||   |       ||  |  | |  ||   |||    |||  | ||||||
Db  112  ESEDEKNQISVK---IKGRAANDPHFPD--SRXHYYENGKDVYDATLNQTEISQNNNKXY  166

Qy  120  LIQLLEDDAQRNFSVWMRWGRVGKMGQHSLVAC-SGNLNKAKEIFQKKFLDKTKNNWEDR  178
         :|||||||: ||| |||| | ||    ||||   ||||||||: |  | ||:|||||||
Db  167  IIQLLEADDGSSYSVWNRWGREGLKGQSSRKDFGKGGLNQAISLFCSKFYEKTKNTFTDR  226

Qy  179  EKFEKVPGKYDMLQMDYATNTQDE---EETKKEESLKSPL-KPESQLDRVQELIKLIC  233
         :| ||  ||||||:|||:|::|    | ||||  |:   ||||  |||||:|||:|||
Db  227  ANFKKVAGKYDMIELDYSTDSKPKNGASTTATTTTKKVVEHKKECSLDERVQELVKLIF  286

Qy  234  NVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFY  293
         |||| || || | ||   ||| | ||||||| | |||||  |||||||||  | | ||
Db  287  DVKMMERTMTEAKYDLKKKMPLGKLSKNQITKGYLVLKQIED-VMGGKSGESLSTLSSRFY  345

Qy  294  TRIPHDFGLRTPPLIRTOKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLH  353
         || |    ||| ||||| ||| : |||| | |||   |  :  ||    |  :| | |
Db  346  TIIPHAFGMSVPPVTNTNQMLIEKMNMLQNLADIEIATNIIKDSESDESNILELHYAKLK  405

Qy  354  CALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLH--NRM  411
          ||||  |:||: : |||  :|||||:  ||   |||:|:||:|||| |||| |
Db  406  TDIQPLDENSCEYKNILLVVKNTY---QGGKKPTIVNIFKIDRDGEADRYKTKKHLGNRK  462

Qy  412  LLWHGSRMSNVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGL  471
         |||||:  : |  :| ||| ||   |  :| || ||||||  ||  ::
Db  463  LLWHGSRLTNYASIISQGLRIAPPEAPVSSYRFGKGVYFADCMSLSANYC-----RTVGF  517

Qy  472  L-----LLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPS-SAHFVTLNGSTVPL  525
               || :|||   :|| |:||:: :|: |  | | ||| |
Db  518  LIDFCMLLGDVALGKTADLAR-DTYMEKPQPNSHSTWALGTVEPDPKVFFQDTEGCITPY  576

Qy  526  G---PASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQ  563
         |    |:  |: |  : ::|||||  | ||:|||:
Db  577  GQMIPSQHKGV-----SCYEHQYVVYDVAQVHLKYLLQLR  611
```

Right column:

```
RP    SEQUENCE FROM N.A.
RX    MEDLINE=96007847; PubMed=7578427;
RA    Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA    Kofler B., Schweiger M., Wagner E.F.;
RT    "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT    ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT    inactivation of the ADPRT gene in the mouse.";
RL    Biochimie 77:444-449(1995).
DR    HSSP; P26446; 1A26.
DR    InterPro; IPR001357; BRCT.
DR    InterPro; IPR001290; PARP.
DR    InterPro; IPR004102; PARP_reg.
DR    Pfam; PF00533; BRCT; 1.
DR    Pfam; PF00644; PARP; 1.
DR    Pfam; PF02877; PARP_reg; 1.
DR    PROSITE; PS50172; BRCT; 1.
SQ    SEQUENCE  607 AA;  67496 MW;  54CDEBBE22079886 CRC64;

Query Match            32.7%;   Score 981.5;  DB 13;  Length 607;
Best Local Similarity  40.9%;   Pred. No. 2.1e-60;
Matches  216;  Conservative 101;  Mismatches 180;  Indels 31;  Gaps 13;

Qy   51  PVAG---GKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLN  107
             |:  | |      |||     ||||||  |||||  | |  ||||||| ||||||||
Db   93  PVAGKSSGKVKEEKGSNKSEKKMK-LTVKGGAAIDPDSELE-DSCHVLETGGKIFSATLG  150

Qy  108  QTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGK-MGQHSLVACSGNLNKAKEIFQKK  166
         : :|       ||:||||:|||||||||:||| |||  :| :|||  ::||||| |
Db  151  LVDITRGTNSYYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSK-EDAIEHFLNL  209

Qy  167  FLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLRVQ  226
         || |||||  ::   | |     |   |     : :  |:    :  ||:|:|  || |
Db  210  YQDKTGNAWHS-PNFTKYPKKFYPLEIDYG---QEEDVVKK---LSVGAGTKSKLAKPVQ  262

Qy  227  ELIKLICNVQAMEEMMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALM  286
         |||||   :: ||:|  |: |  ||||||:|:|||||| ||:|||    |||||||| |
Db  263  ELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAVSESLSEARLL  322

Qy  287  EACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPE-HPL  345
           | |   |||  ||  ||||||||||| ::||| | |||  |  |:| | |  |
Db  323  DLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVAYSLLRGGADDGEKDPI  382

Qy  346  DQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR-  404
         |:|  |   | :|  |  | :|: :| ::|||:||||||:|:||:||:|||::|||: :
Db  383  DVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNAYDLEVLEIFKIDREGEYQRYKP  442

Qy  405  -EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFA  463
            || ||||||||||||||: ::   |:|||||| | :| || ||||||  ||  ::
Db  443  FKQLNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANYCHA  502

Qy  464  SRLKNTGLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGSTV  523
           | ||| || :|||   |:| |:||:: :|: |  | | ||| |
Db  503  MPGSPIGLILLGEVALGNMHE-LKAASQITKLPKGKHSVKGLGRTAPDPSATVQLDGVDV  561

Qy  524  PLGPA-----SDTGILNPDGYTLNYINEYIVYNPNQVRMRYLLKVQFNF  566
          | |       | : | |  : ::|||||  | ||:|||:
Db  562  PLGKGTSANISDTSLL--------YNEYIVYDIAQVNLKYLLKLKFNY  601

RESULT  7
Q9TX05
ID    Q9TX05      PRELIMINARY;     PRT;     593 AA.
AC    Q9TX05;
DT    01-MAY-2000 (TrEMBLrel. 13, Created)
DT    01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT    01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE    NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS    Drosophila sp. (Fruit fly).
OC    Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC    Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC    Ephydroidea; Drosophilidae; Drosophila.
OX    NCBI_TaxID=7242;
```

Lower left column:

```
RESULT  6
Q9PS81
ID    Q9PS81      PRELIMINARY;     PRT;     607 AA.
AC    Q9PS81;
DT    01-MAY-2000 (TrEMBLrel. 13, Created)
DT    01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT    01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE    NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS    Xenopus.
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC    Xenopodinae; Xenopus.
OX    NCBI_TaxID=8353;
RN    [1]
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   593 AA;  68018 MW;  D9BA37E3B8E7CCD CRC64;

Query Match             32.4%;  Score 972.5;  DB 5;  Length 593;
Best Local Similarity   38.3%;  Pred. No. 8.6e-60;
Matches  215;  Conservative 110;  Mismatches 199;  Indels 37;  Gaps 11;

Qy      12 GRARALNESKRVNNGNTAPEDSSPAKKTRRCCQRQE--SKKMPVAGGKANKDRTEDKQDES 69
              ::  :::|.|||: ||:|: | : : :|::|       |:::
Db      62 GAIKYIN-STCICSWGTDPKSRIPKETTKSLNSNSIYTKSMPVS-----------:YTKSMPVS 104

Qy      70 VKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQ 129
             ||  :  ::|:|: |:|:|||||:|||||||||||||| ||||||||||||:|||||||||
Db     105 -RTFKVKDGLAVDPD-SGLEDIAHVYVDSNNKYSVVLGLTDIQRNKNSYYKVQLLADKK 162

Qy     130 RNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPGKYD 189
             |:||:|:||||||||||||||||||||||||:|:||||||||:||||||||| || : |
Db     163 EKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVKRTGRMY 222

Qy     190 MLQMDYATNTQDEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEMMEMKYNT 249
             :|:|: | |: :||||||||||||:|| :: :||||||||||||||||||| :||
Db     223 PIEIQY---DDDQKLVKHESHFF---TSKLEISVQNLIKLIFDIDSMNKTLMEFHIDM 274

Qy     250 KKAPLGKLTVAQIKAGYQSLKKIEDCIRAGQHGRALMEACNEFYTRIPHDFGLRTPPLIR 309
             ||:|||||:|||||||:: :|: :|| ::|:|||||||:||||:||| ||:|||:| |
Db     275 DKMPLGKLSAHQIQSAYRVVKEIYNVLECGSNTAKLIDATNRFYTLIPHNFGVQLPTLIE 334

Qy     310 TQKELSEKIQLLEALGDIEIAIKLVKTE-LQSPEHPLDQHYRNLHCALRPLDHESYEFKV 368
             ||||||:||||||||||||||||:||:| |||||||||||||||||| |||||||||||
Db     335 THQQIEDLRQMLDSLAEIEVAYSIIKSEDVSDACNPLDNHYAQIKTQLVALDKNSEEFSI 394

Qy     369 ISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE--KEAFREDLHNRMLLWHGSRMSNWVGIL 426
             |:||:||||||||||||| ::|| |: :|: :||: | :|||||||||||||: ||||
Db     395 LSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARREKPFKKLHNRKLLWHGSRLTNFVGIL 454

Qy     427 SHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELL 486
             |||||||||||:|||||||||||||||||||||||||:|||||||||:||||||:||:||
Db     455 SHGLRIAPPEAPPTGYMFGKGIYFADMVSKSANYCCTSQQNSTGLMLLSEVALGDMMECT 514

Qy     487 EANPKAEGLLQGKHSTKGLGKMAPS-SAHFVTLNGSTVPLGPASDTGILNPDGYTLNYNE 545
             || |||:| :|| ||||||||||| ||:||| |||:: |||| | |:||||:||||||
Db     515 SAK-YINKLSNNKHSCFGRGRTMPDDPTKSYIRSDGVEIPYGETITDEHLKS---SLLYNE 570

Qy     546 YIVYNPNQVRMRYLLKVQFNE 566
             |||| ||| | ||||||:: :
Db     571 YIVYDVAQVNIQYLFRMEFKY 591
```

RESULT 8

```
GN   PARP-1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LANDSBERG ERECTA;
RA   Doucet-chabeaud G., Kazmaier M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ131705; CAA10482.1; -.
DR   HSSP; P26446; 1A26.
KW   Transferase; Glycosyltransferase; NAD.
FT   CHAIN      2     983       POLY(ADP-RIBOSE) POLYMERASE.
SQ   SEQUENCE   983 AA;  111232 MW;  468E12A8EF1B6F4F CRC64;

Query Match             32.2%;  Score 966;  DB 10;  Length 983;
Best Local Similarity   38.8%;  Pred. No. 4.8e-59;
Matches  213;  Conservative 115;  Mismatches 151;  Indels 70;  Gaps 17;

Qy      45 QESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAPVDPECTAKVGKAHVYCEGNDVYDV 104
             :: |:: ::|  |::|
Db     472 KKQRKLPF-----DKYKIED-TSESLVTVKVKGRSAVH-EASGLQEHCHILEDGNSIYNT 524

Qy     105 MLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVG--KMGQHSLVACSGNLNKAKEI 162
             ::||::|||:|||||:|:||||:: ||: : || :||||  |: |||||||||||||
Db     525 TLSMSDLSTGINSYYIIQIIQEDKGSDCYVFRKWGRVGNEKIGG-------NKVEEM 574

Qy     163 -------FQKKFLDKTKNNWEDREK---FEKVPGKYDMLQMDYATNTQDEEETKKEESL 211
                    | :|| :: ::||    |::|| |:|::    : :|
Db     575 SKSDAVHEFKRLFLEKTGNTWESWEQKTNFQKQPGKFLPLDIDYGVN-------KQVAK 626

Qy     212 KSPLKPESQLDLRVQELIKLICNVQAMEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKK 271
             ||: | || || |||||||||:|||||||| ||:|:| |||||:| :||||:|:|::
Db     627 KEPFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTE 686

Qy     272 IEDCI----RAGQHGRALMEACNEFYTRIP-HDFGLRTPPLIRTQKELSEKIQLLEAL 324
             |::     ||::|::| ||:|::|||:||  :|: :|: |::| :|||: :|::||||
Db     687 IQRLLTESDPQPTMKESLLVDASNRFFTMIPSIH-----PHIIRDEDDFKSKVKMLEAL 740

Qy     325 GDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDY 384
             |||||::|||:|:|||||||||||||| || |||||||||:||||||::|||||||||
Db     741 QDIEIASRIVGFDVDSTE-SLDDKYKKLHCDISPLPHDSEDYRLIEKYLNTTHAPTHTEW 799

Qy     385 TMTLLDLFEVEKDGEKEAF---REDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITG 441
             |:|||| |: |:|| || |    |:||||||||||||||||||||:|||||||||||||
Db     800 SLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPEAPATG 859

Qy     442 YMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVALGQCNELLEAN-----PKAEGLL 496
             ||||||||||||:|||||:||||||:|:|| ::|||||: :||||
Db     860 YMFGKGIYFADLVSKSAQYCYTCKKNPVGLMLLSEVALGEIHELTKAKYMDKPPR----- 914

Qy     497 QGKHSTKGLGKMAPSSAHFVTLNGS-TVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVR 555
             |||||||||||||| :||||| ||  |:| ||| | |:||||:||||| : |::  |
Db     915 -GKHSTKGLGKKVPQDSEFAKWRGDVTVPCGKPVSSKVKASE---LMYNEYIVYDTAQVK 970

Qy     556 MRYLLKVQF 564
             ::|||| :|
Db     971 LQFLLKVRF 979
```

RESULT 9

```
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.;
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.;
RC   STRAIN=CV. COLUMBIA;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006593; AAD20677.1; -.
HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001510; znf-PARP.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE 1009 AA;  114133 MW;  CDE6E41CC2A3A2DB CRC64;

Query Match             32.2%;  Score 966;  DB 10;  Length 1009;
Best Local Similarity   38.8%;  Pred. No. 5e-59;
Matches  213;  Conservative  115;  Mismatches  151;  Indels  70;  Gaps  17;

Qy         45 QESKKMPVAGGKANKDRTEDKQDESVKALILKGKAPVDPECTAKVGKAHVYCEGNDVYDV 104
              :: :|:|              || ||  ::|||::  :||::|   :|: :||:|:
Db        498 KKQRKLPF-----DKYKIED-TSESLVTVKVKGRSAVH-EASGLQEHCHILEDGNSIYNT 550

Qy        105 MLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVG--KMGQHSLVACSGNLNKAKEI 162
              |:  ||||:|||:|          ||:|| || :||| ||
Db        551 TLSMSDLSTGINSYYILQIIQEDKGSDCYYFRKWGRVGNEKIGG--------NKVEEM 600

Qy        163 -------FQKKFLDKTKNNWEDREK--FEKVPGKYDMLQMDYATNTQDEEETKKEESL 211
                     |:|||||||||| ||:| |||:|  | |:|  ||| :|||:::||:||
Db        601 SKSDAVHEFKRLFLEKTGNTWESWEQKTNFQKQPGKFLPLDIDYGVN------KQVAK 652

Qy        212 KSPLKPESQLDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKK 271
              | ||| | |::|||:||||:|||| |::|||||:|| :  |:  ||| ::| |:
Db        653 KEPFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTE 712

Qy        272 IEDCI-----RAGQHGRALMEACNEFYTRIP--HDFGLRTPPLIRTQKELSEKIQLLEAL 324
              :|||:      |:| | |:  ||  ||| :||   |:  |: :|  ::|:||::|:|||
Db        713 IQRLLTESDPQPTMKESLLVDASNRFFTMIPSIH-----PHIIRDEDFKSKVKMLEAL 766

Qy        325 GDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDY 384
              |||||||||:|||| :| : |||| ::|:| | |||:| :::| ||:|:|||  |
Db        767 QDIEIASRIVGFDVDSTE-SLDDKYKKLHCDISPLPHDSEDYRLIEKYLNTTHAPTHTEW 825

Qy        385 TMTLLDLFEVKDGEKEAF--REDLHNRMLLWHGSRMSNWVGIISHGLRIAPPEAPITG 441
              |:  |:::||:||      :|||  ||:||||||| |||| |||:||||||||||| |
Db        826 SLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGIINQGLRIAPPEAPATG 885

Qy        442 YMFGKGIYFADMSSKSANYCFASRLKNTGLLLSEVALGQCNELLEAN----PKAEGLL 496
              ||||||||| |::|||:|   || |  |||||||| ||  :|: |||    |||| ||
Db        886 YMFGKGIYFADLVSKSAQYCTCKKNPVGLMLLSEVALGEIHELRKAKYMDKPPR---- 940

Qy        497 QGKHSTKGLGKMAPSSAHFVTLNGS-TVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVR 555
              ||||| ||||| : : ||| | |    ||||| || ||  ||  :  | ::
Db        941 -GKHSTKGLGKKVPQDSEFAKWRGDVTVPCGKPVSSKVKASE--LMYNEYIVYDTAQVK 996

Qy        556 MRYLLKVQF 564
              |:||||:|
Db        997 LQFLLKVRF 1005
```

```
RESULT 10
024570
ID   024570;                       PRELIMINARY;       PRT;       969 AA.
AC   024570;
DT   01-JAN-1998 (TREMBLrel. 05, Created)
DT   01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babiychuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants possess two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ225589; CAA10889.1; -.
HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE 969 AA;  109128 MW;  EB23AC62EEC14009 CRC64;

Query Match             32.1%;  Score 961.5;  DB 10;  Length 969;
Best Local Similarity   39.3%;  Pred. NO. 9.8e-59;
Matches  223;  Conservative  104;  Mismatches  179;  Indels  61;  Gaps  18;

Qy         32 DSSPAKKTRRCQRQESKKMPVAGG------KANK-------DRTEDKQDESVKALLL 75
              :::  :  |  | | ||::||| ::      :|:|       |:|::||||:|||
Db        428 ENAEVKKARRL------KIPIVREGYIGECVKKNKMLPFDLYKLENALESSKGSTVVKV 481

Qy         76 KGKAPVDPECTAKVGKAHVYCEGNDYDVDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVW 135
              |||| | |  |  || |: :||  ||:| |:| :|| |:|||:| | |:::|
Db        482 KGRSAVH-ESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQIIEQDDGSECYF 540

Qy        136 MRWGRVG--KMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE--KPEKVPGKYDM 190
              ||||| | |||:||||| ::  |:|:|| ::|| ||||||||  ||  |  :||| ||
Db        541 RKWGRVGSEKIGGQKLEEMS--KTEAIKEFKRLFLEKTGNSWEAWECKTNFRKQPGRFYP 598

Qy        191 LQMDYATNTQDEEETKKEESLKSPLKPESQLDRVQELIKLICNVQAMEEMMEMKYNTK 250
              | |:   |:||:|||::|||||: ||| |  :   |::  |||| |:::|| :||:|
Db        599 LDVDYGVKKAPKRKDISE--MKSSLAPQ------LLELMKMLFNVETYRAAMMEFEINMS 650

Qy        251 KAPLGKLTVAQIKAGYQSLKKIEDCIR-AGQHGRALME-----ACNEFYTRIP-HDFGL 302
              |||||::| |:|| |:|| :|::|  | | :| | |:|     |  :||| :|  ||:
Db        651 EMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAASNRFFTLIPSIH---- 706
```

```
Qy  136 MRWGRVG--KMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDRE---KFEKVPGKYDM 190
            :|||| |     | :  ||     : ||  |  ||   :|:|| |   || ::| ||
Db  552 RKWGRVGSEKIGGQKLEEMS--KTEAIKEFKRLFLEKTGNSWEAWECKTNFRKQPGRFYP 609

Qy  191 LQMDYAYNTQDEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEMMEMKYNTK 250
            | :|   | |   | |  :  :    || |    || |:| :
Db  610 LDVDYGVKKAPKRKDISE--MKSSLAPQ-----LLELMKMLFNVETYRAAMMEFEINMS 661

Qy  251 KAPLGKLTVAQIKAGYQSLKKIEDCIR-AGQHGRALME---ACNEFYTRIP-HDFGL 302
        . |:||| :| ||    || :| :| | :: |  | :    |||| || |   ||::|
Db  662 EMPLGKLSKENIEKGFEALTEIQNLLKDTADQALAVRESLIVAASNRFFTLIPSIH---- 717

Qy  303 RTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHE 362
          | :|| :   | |:| |||| |||||| :|  |  | : | |  | ::| || || || 
Db  718 --PHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDE-SLDDKYMKLHCDITPLAHD 774

Qy  363 SYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE---KEAFREDLHNRMLLWHGSRM 419
         ::| |: :|| ||||||| |: |||: | | ||||     | ||:|||:||||||||| 
Db  775 SEDYKLIEQYLLNTHAPTHKDWSLELEEVFSLDRDGELNKYSRYKNNLHNKMLLWHGSRL 834

Qy  420 SNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVAL 479
        :| |||||:|||||||||| :||||||| ||||: |||| |:  :   :|| |:||||||
Db  835 TNFVGILSQGLRIAPPEAPVTGYMFGKGLYFADLVSKSAQYCYVDRNNPVGLMLLSEVAL 894

Qy  480 GQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGS-TVPLGPASDTGILNPDG 538
        | : || :|:   | ||||||||||||| ||: :  :  | | |||  :   ::|    :
Db  895 GDMYELKKAT-SMDKPPRGKHSTKGLGKTVPLESEFVKWRDDVVVPCGKPVPSSIRSSE- 952

Qy  539 YTLNYNEYIVYNPNQVRMRYLLKVQFN 565
           |:|||||||||:::|:|: | |:
Db  953 --LMYNEYIVYNTSQVKMQFLLKVRFH 977


RESULT 12
Q9N4H4
ID   Q9N4H4         PRELIMINARY;         PRT;   945 AA.
AC   Q9N4H4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 108.0 KDA PROTEIN.
GN   Y71F9AL.18.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw-Cordum H., Scott K., Graves T.;
RT   "The sequence of C. elegans cosmid Y71F9AL.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024200; AAF36011.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 1.
```

```
Qy  303 RTPPLIRTQKELSEKIQLLEALGDIEIAIKLVKTELQSPEHPLDQHYRNLHCALRPLDHE 362
          | :|| :   | |:| |||| |||||| :|  |  | : | |  | ::| || || || 
Db  707 --PHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSDSDE-SLDDKYMKLHCDITPLAHD 763

Qy  363 SYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEKDGE---KEAFREDLHNRMLLWHGSRM 419
         ::| |: :|| ||||||| |: |||: | | ||||     | ||:|||:||||||||| 
Db  764 SEDYKLIEQYLLNTHAPTHKDWSLELEEVFSLDRDGELNKYSRYKNNLHNKMLLWHGSRL 823

Qy  420 SNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSEVAL 479
        :| |||||:|||||||||| :||||||| ||||: |||| |:  :   :|| |:||||||
Db  824 TNFVGLISQGLRIAPPEAPVTGYMFGKGLYFADLVSKSAQYCYVDRNNPVGLMLLSEVAL 883

Qy  480 GQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLNGS-TVPLGPASDTGILNPDG 538
        | : || :|:   | ||||||||||||| ||: :  :  | | |||  :   ::|    :
Db  884 GDMYELKKAT-SMDKPPRGKHSTKGLGKTVPLESEFVKWRDDVVVPCGKPVPSSIRSSE- 941

Qy  539 YTLNYNEYIVYNPNQVRMRYLLKVQFN 565
           |:|||||||||:::|:|: | |:
Db  942 --LMYNEYIVYNTSQVKMQFLLKVRFH 966


RESULT 11
Q9ZSV1
ID   Q9ZSV1         PRELIMINARY;         PRT;   980 AA.
AC   Q9ZSV1;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP)-RIBOSE POLYMERASE (EC 2.4.2.30).
GN   PARP1.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99026291; PubMed=9808734;
RA   Mahajan P.B., Zuo Z.;
RT   "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL   Plant Physiol. 118:895-905(1998).
DR   EMBL; AF093627; AAC79704.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 2.
KW   Transferase; Glycosyltransferase.
SQ   SEQUENCE 980 AA; 110475 MW; 9D8AED26BC37E5C1 CRC64;

Query Match             32.1%;  Score 961.5;  DB 10;  Length 980;
Best Local Similarity   39.38%;  Pred. No. 9.9e-59;
Matches  223;  Conservative  104;  Mismatches 179;  Indels  61;  Gaps  18;

Qy   32 DSSPAKKTRRCQRQESKKMPVAGG------KANK------DRTEDKQDESVKALLL 75
            ::|  |:| :: :   :  :          ||        | :||  : ||  :
Db  439 ENAEVRKARRL-----KIPIVREGYIGECVKKNKMLPFDLYKLENALESSKGSTVTVKV 492

Qy   76 KGKAPVDPECTAKVGKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVW 135
        || :|| |  |  ||||||  || |||:: ||: ||  :|: |||:||:|:|  |  : :
Db  493 KGRSAVH-ESSGLQDTAHILEDGKSIYNATLNMSDLALGVNSYYVLQIIEQDDGSECYVF 551
```

Left column:

```
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   945 AA;  108006 MW;  1D0A62C954BC6AD9 CRC64;

Query Match             25.9%;  Score 777.5;  DB 5;  Length 945;
Best Local Similarity   34.7%;  Pred. No. 6.8e-46;
Matches  196;  Conservative 103;  Mismatches 213;  Indels  53;  Gaps  17;

Qy    21 KRVNGNTAPEDSSPAKKTRRCQPQESKKMPVAGGKANKDRTEDKQDESVKALLLKGKAP 80
Db   409 KRKSKGGTRGEQFIYAAEA---FDSTNNVPIKVGDLTSTNTH------IIKKGTV 454

Qy    81 VDPECTAKVGKAHVYCEGND--VYDVMLNQTNLQFNNKYYLIQLLEDDAQRNFSVMRWRW 138
Db   455 VDAK-FALADRCHVFKNEIDGSLYQATLSFTDLTQNKNSYYKIQLLKDDQRENYYVFRSW 513

Qy   139 GRVG-KMGQHSLVACSGNLNKAKEIFQKFLDKTKNNWEDREKFEKVPGKYDMLQMDYAT 197
Db   514 GRVGTEVGGNKHESYS-NSNEAILKFQDVFHEKTKNDWIYRKIHFRKMPGMFSYVETDYSE 572

Qy   198 NTQ--DEEETKKEESLKSPLKPESQLDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLG 255
Db   573 FAQTTDTEITPGSKTL---------LPKSVKEVMSIFDVENMKSALKSFEMDVNKMPLG 623

Qy   256 KLTVAQIKAGYQSLKKIED-CIRAGQHGRALMEACNEFVTRIPHDFGLRTPPLIRTQKEL 314
Db   624 RLSHNQINLAFEVLNDISDLLVKLPIDASRILDFSNKFYTIIPHNFGMRVPEPIDSFHKI 683

Qy   315 SEKIQLLEALGDIEIAIKLV------KTELQSPEHPLDQHYRNLHCALRPLDHESYEFK 367
Db   684 KEKNNMLNALLDIKFAYDQISGGDVPASTSLGI--DPVDINYQKLKCIMEPLQQGCDDWN 741

Qy   368 VISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFREDLHNRMLLWHGSRMSNWVGILS 427
Db   742 MIHQYLKNTHGATH-DLKVELIDILKLNRDNESSKFKRHIGNRLLWHGSGKMNFAGILG 800

Qy   428 HGLRIAPPEAPITGYMFGKGIYFADMSSKSSANYCFASRLKNTGLLLLSEVALGQCNELLE 487
Db   801 QGLRIAPPEAPVSGYMFGKGVYFADMFSKSFFYCRAN-AKEEAYLLLCDVALGNVQQLMA 859

Qy   488 A-NPKAEGLLQGKHSTKGLGKMAPSS-AHFVTLNGSTVPLGPASDTGILNPDG----YT 540
Db   860 SKNVSRQTLPAGFQSVQGLGRQCPREIGSYNHPDGYTIPLG----LTYMQLQGKQDVDYH 915

Qy   541 LNYNEYIVYNPNQVRMRYLLKVQFN 565
Db   916 LLYNEFIVYDVDQIQLKYLVRVKMH 940

RESULT 13
ID   Q96CG2      PRELIMINARY;      PRT;   533 AA.
AC   Q96CG2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 60.1 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014260; AAH14260.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   533 AA;  60089 MW;  6296A0E439CC7767 CRC64;

Query Match             24.2%;  Score 725;  DB 4;  Length 533;
```

Right column:

```
Best Local Similarity   35.0%;  Pred. No. 1.5e-42;
Matches  194;  Conservative  97;  Mismatches 188;  Indels  76;  Gaps  20;

Qy    46 ESKKMPVAGGKANKDRTED---KQDESVKALLLKGK-APVDPECTAKVGK-AHVYCEGND 100
Db    15 EKKKGRQAG--------REEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVY---E 64

Qy   101 VYDVMLNQTNLQFNNKYYLIQLLEDDAQRNFSVMRWRGVRGVKMGQHSLVACSGNLNKAK 160
Db    65 DYNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWRGVEVGQ-SKINHFTRLEDAK 122

Qy   161 EIFQKKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ 220
Db   123 KDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV--QAEDEAQEAVVKDRGPVRTVTK 178

Qy   221 -----LDLRVQELIKLICNVQAMEMMMEMKYNTKKAPLGKLITVAQIKAGYQSLKKIED 274
Db   179 RVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEE 238

Qy   275 CIRA-GQHGRALMEACNEFVTRIPHDFGLRITPPLIRTQKELSEKIQLLEALGDIEI--AI 331
Db   239 ALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQAL 298

Qy   332 KLVKTELQSPE--HPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTL 388
Db   299 QAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS--NHRCPTL 355

Qy   389 LDLFEVEKDGEKEAFR---EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGK 446
Db   356 QHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGK 411

Qy   447 GIYFADMSSKSSANYCFASR--LKNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKG 504
Db   412 GIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL------------ 459

Qy   505 LGKMAPSSAHFVTLNGSTVPLGPASDT------GILNPDGY------TLNYNEYI 547
Db   460 --KSPPGFGDSVIARGHTEP-DPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYL 516

Qy   548 VYNPNQVRMRYLLKV 562
Db   517 IYQESQCRLRYLLEV 531

RESULT 14
ID   Q91YR6      PRELIMINARY;      PRT;   528 AA.
AC   Q91YR6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 59.4 KDA PROTEIN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014870; AAH14870.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   528 AA;  59413 MW;  82EFB0C498EB5F74 CRC64;

Query Match             23.2%;  Score 696.5;  DB 11;  Length 528;
Best Local Similarity   34.5%;  Pred. No. 1.4e-40;
Matches  184;  Conservative  92;  Mismatches 213;  Indels  45;  Gaps  17;

Qy    54 GGKANKDRTEDKQDESVKALLLKGKAP------VDPECTAKVGKAHVYCEGNDVYDVML 106
Db    13 GSKKQRQGTEEEDSFRSTAEALRA-APADNRVIRVDPSCPFSRNPG---IQVHEDYDCTL 68
```

```
Qy   19 ESKRVNNGNTAPEDSSPAKKTRRCQRQESKKMPVAGGKAANKDRTEDKQDESVKALLLKGK  78
Db  189 EALAAKGGSTEPATPASASPT-----PPEAETPVLSAEGSPE-SSNKRPASAEIIEIDGE  242

Qy   79 APVDPECTAK--------VGKAHVYCEGNDYDVMLNQTNLQFNNNKYYLIQLLED        126
Db  243 GNPDENDFAKKRRMKKEARLMEVQKKRMKKQSDLLWEYRQIFERMPYTDN----ISILRE  298

Qy  127 DAQRNFSVWMRWGRVGKMGQHSLVACSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPG  186
Db  299 NEQD--------IPEGHDRTAQDVFH----EKTKNDWIYRKHFRKMPG              334

Qy  187 KYDMLQMDYATNTQDEEETKKEESLKSPLKPESQ-LDLRVQELIKLICNVQAMEMMME    244
Db  335 MFSYVETDYSEFVGTNNGHKKKITPGSKTLLPKSVKEVVMSIFDVENMKSALKS        394

Qy  245 MKYNTKKAPLGKLTVAQIKAGYQSLKKIED-CIRAGQHGRALMEACNEFYTRIPHDFGLR  303
Db  395 FEIDVNKMPLGRLSHNQINLAFEVLNDISDLLVKLPIDASKILDFSNKFYTIIPHNFGMR  454

Qy  304 TPPLIRTQKELSEKIQLLEALGDIEIAIKLV------KTELQSPEHPLDQHYRNLHCAL   356
Db  455 VPEPIDSFHKIKEKNNMLNALLDIKFAYDQISGGDVPASTSLSI-DPVDINYRKLKCIM   512

Qy  357 RPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVERDGEKEAFREDLHNRMLLWHG  416
Db  513 EPLQQGCDDWNMIHQYLKNTHGATH-DLKVELIDILKVNRDNESSKFKRHIGNRRLLWHG  571

Qy  417 SRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMSSKSANYCFASRLKNTGLLLLSE  476
Db  572 SGKMNFAGILGQGLRIAPPEAPVSGYMFGKGVYFADMFSKSFFYCRAN-AKEEAYLLCD   630

Qy  477 VALGQCNELLEA-NPKAEGLLQGKHSTKGLGKMAPSS-AHFVTLNGSTVPLGPA--SDTG  532
Db  631 VALGNVQQLMASKNVSRQTLPAGFQSVGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQG   690

Qy  533 ILNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFN                            565
Db  691 KQNVD-YHLLYNEFIVYDVDQIQLKLIVRVMH                             722
```

```
Qy  107 NQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSL--VACSGNLNKAKEIFQ  164
Db   69 NQNIGNNNNKFYIIQLLEEGSR--FFCWNRWGRVGEVGQSKMNHFTC---LEDAKKDFK   123

Qy  165 KKFLDKTKNNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEESLKSPLKPESQLDLR  224
Db  124 KKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTVVKPCS-LDPA  182

Qy  225 VQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIKAGYQSLKKIEDCIR--AGQHG   282
Db  183 TQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAMKNPTGD-G  241

Qy  283 RALMEACNEFYTRIPHDFGLRTPPLIRTQKELSEKIQLLEALGDIEIAIKL------VKT  336
Db  242 QSLEELSSCFYIVIPHNFGRSRPPINSPDVLQAKKDMLLVLADIELVQTLQAAPGEEEE   301

Qy  337 ELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQSTHAPTHSDYTMTLLDLFEVEK  396
Db  302 KVEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQT---GNSYRCPNLRHVWKVNR   358

Qy  397 DGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIYFADMS  454
Db  359 EGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVGKGIYFASEN  414

Qy  455 SKSANYCFASRL-KNTGLLLLSEVALGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSS  512
Db  415 SKSAGYVTTMHCCGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP  474

Qy  513 AHFV--TLNGS--TVPLGPASDTGILNPDGYTLNYNEYIVYNPNQVRMRYLLKV        562
Db  475 AQDIELELDGQPVVVPQGGPPVQCPSFKSSSF--SQSEYLIYKESQCRLRYLLEI       526
```

```
RESULT 15
Q9XUA5
ID   Q9XUA5      PRELIMINARY;    PRT;    727 AA.
AC   Q9XUA5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC8.1 PROTEIN.
GN   AC8.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McMurray A.A.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; 283097; CAB05448.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 1.
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 1.
SQ   SEQUENCE 727 AA; 82884 MW; 530ABA8E991FFEFD CRC64;
```

```
Query Match       22.9%;    Score 685.5;  DB 5;  Length 727;
Best Local Similarity   31.1%;    Pred. No. 1.3e-39;
Matches 178;  Conservative 110;  Mismatches 220;  Indels 65;  Gaps 16;
```

Search completed: August 29, 2002, 08:01:28
Job time: 360 sec

THIS PAGE BLANK (USPTO)

```
                   GenCore version 4.5
         Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 29, 2002, 07:57:12 ; Search time 124.84 Seconds
                                             (without alignments)
                                             474.226 Million cell updates/sec

Title:           US-09-701-586B-4
Perfect score:   2823
Sequence:        1 MAPKPKPWVQTEGPEKKKGR.........EYLIYQESQCRLRYLLEVHL 533

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters:       747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_032802:*
   1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
   2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
   3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
   4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
   5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
   6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
   7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
   8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
   9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
  10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
  11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
  12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
  13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
  14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
  15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
  16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
  17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
  18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
  19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
  20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
  21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
  22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2823	100.0	533	21	AAY51175	Human brain PARP3
2	2823	100.0	540	21	AAY51176	Human uterus type
3	2811	99.6	533	22	AAU29021	Human PARP-3 prote
4	2253.5	99.8	528	21	AAY51178	Murine PARP1 (shor
5	2241	79.4	533	21	AAY51177	Murine PARP1 (long
6	738	26.1	522	22	AAU29022	Mouse PARP-2 prote
7	729.5	25.8	583	21	AAB47029	hPARP2, Homo sapi
8	726.5	25.7	534	21	AAB42909	Human ORFX ORF2673
9	726.5	25.7	534	22	AAU29023	Human PARP-2 prote
10	725.5	25.7	534	22	AAU29020	Human PARP-2 prote
11	725	25.7	570	21	AAY51174	Human brain PARP2
12	725	25.7	570	22	AAB11480	Human brain poly-A
13	722	25.6	521	22	AAB60693	Human poly(ADP-rib
14	688.5	24.4	637	21	AAY68835	The poly(ADP-ribos
15	682.5	24.2	531	22	AAB93513	Human protein sequ
16	670.5	23.8	653	21	AAY68834	A poly(ADP-ribose)
17	668.5	23.7	1014	21	AAY58043	Human poly (ADP-ri
18	668.5	23.7	1014	22	AAU29019	Human PARP-1 prote
19	668.5	23.7	1014	22	AAB66296	Human tankyrase2 r
20	667.5	23.6	1014	21	AAY49939	Human nuclear NAD+
21	666.5	23.6	1013	17	AAR99642	Poly(ADP-ribose) p
22	660.5	23.4	1014	20	AAY33699	Human poly(ADP-rib
23	638	22.6	557	22	ABB66431	Drosophila melanog
24	637	22.6	1063	21	AAB47032	Fusion protein PAR
25	578	20.5	969	21	AAY68833	A poly(ADP-ribose)
26	578	20.5	980	21	AAY68839	Maize poly ADP-rib
27	475	16.8	982	20	AAY28464	Novel human neopla
28	472.5	16.7	379	22	AAU21687	N-terminal fragmen
29	447	15.8	360	22	AAB47030	Fusion protein of
30	410.5	14.5	1010	21	AAY68840	Human DNA repair a
31	394	14.0	294	22	AAU20129	Novel human neopla
32	394	14.0	294	22	AAU21810	C-terminal fragmen
33	390	13.8	287	22	AAB47031	Human tankyrase2 e
34	284.5	10.1	1099	22	AAB66301	Human DNA repair a
35	267.5	9.5	227	22	AAU20130	Novel human neopla
36	267.5	9.5	227	22	AAU21811	Novel human neopla
37	258	9.1	1730	22	AAU33242	Novel human secret
38	256	9.1	1724	22	AAY54373	cDNA sequence enco
39	256	9.1	1724	22	AAB51022	Human minor vault
40	216.5	7.7	190	22	AAU21688	Novel human neopla
41	138.5	4.9	1327	21	AAB27212	Human tankyrase I
42	138.5	4.9	1327	21	AAY44402	Human tankyrase.
43	138.5	4.9	1327	22	AAB66279	Human tankyrase1 s
44	137	4.9	756	22	AAB66286	Human tankyrase2 c
45	137	4.9	784	22	AAB66285	Human tankyrase2 c

ALIGNMENTS

```
RESULT 1
AAY51175
ID   AAY51175 standard; Protein; 533 AA.
XX
AC   AAY51175;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human brain PARP3 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP3.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;  99WO-EP03889.
XX
PR   05-JUN-1998;  98DE-1025213.
PR   01-MAR-1999;  99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AZ44288.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
```

PT   conditions -
XX
PS   Claim 4; Page 57-59; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP3 protein used in the method of the invention.
XX
SQ   Sequence 533 AA;
XX

Query Match   100.0%;   Score 2823;   DB 21;   Length 533;
Best Local Similarity 100.0%;   Pred. No. 1.1e-249;
Matches 533; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MAPKPKPWVQTEGPEKKGKRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 mapkpkpwvqtegpekkgkrqagreedpfrstaealkaipaekriirvdptcplssnpgt 60

Qy   61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftrled 120

Qy 121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrgpvrtvtkrv 180

Qy 181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPIGKLSKQQIARGFEALEALEEAL 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 qpcsldpatqklitnifskemfkntmalmdldvkkmpigklskqqiargfealealeeal 240

Qy 241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 kgptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladielaqalqa 300

Qy 301 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 vseqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcptlqhiwk 360

Qy 361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 vnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiyfasens 420

Qy 421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 ksagyvigmkcgahhvgymflgevalgrehhintdnpslkspppgfdsviarghtepdpt 480

Qy 481 QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 qdteleldgqqvvvpqgqpvpcpefsssstfsqseyliyqesqcrlryllevhl 533

RESULT 2
AAY51176
ID   AAY51176 standard; Protein; 540 AA.
XX
AC   AAY51176;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Human uterus type 2 PARP3 protein.
XX

KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP3.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB: AA244289.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 62-64; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP3 protein used in the method of the invention.
XX
SQ   Sequence 540 AA;
XX

Query Match   100.0%;   Score 2823;   DB 21;   Length 540;
Best Local Similarity 100.0%;   Pred. No. 1.1e-249;
Matches 533; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MAPKPKPWVQTEGPEKKGKRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8 mapkpkpwvqtegpekkgkrqagreedpfrstaealkaipaekriirvdptcplssnpgt 67

Qy   61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftrled 127

Qy 121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 128 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrgpvrtvtkrv 187

Qy 181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPIGKLSKQQIARGFEALEALEEAL 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 188 qpcsldpatqklitnifskemfkntmalmdldvkkmpigklskqqiargfealealeeal 247

Qy 241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 248 kgptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladielaqalqa 307

Qy 301 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 308 vseqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcptlqhiwk 367

```
Best Local Similarity 99.6%;    Pred. No. 1.3e-248;
Matches  531;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps
                                                   Mismatches   2;            Gaps

Qy     1 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 mapkpkpwvqtegpekkkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt 60

Qy    61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 qvyedynctlnqtniennnkkfyiiqllqdsnrfftcwnrwgrvgevgqskinhftrled 120

Qy   121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRVTKRV 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrapvrvtkrv 180

Qy   181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 qpcsldpatqklitnifskemfkntmalmdldvkkmplgklskqqiargfealealeeal 240

Qy   241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELIQAKKDMLLVLADIELAQALQA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 kgptdggqsleelsshfytviphnfghsqpppinspeliqakkdmllvladielaqalqa 300

Qy   301 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPPTLQHIWK 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 vseqektveevphpldrdyqllkcqlqldsgapeykvigtylegtgsnhrcptlqhiwk 360

Qy   361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 vnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiyfasens 420

Qy   421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ksagyvigmkcgahhvgymflgevalgrehhintdnpslkspppgfdsviarghtepdpt 480

Qy   481 QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 qdteleldgqqvvvpqgqpvpcpefssstfsqseyliyqesqcrlryllevhl 533
```

RESULT 4
```
AAY51178
ID   AAY51178 standard; Protein; 528 AA.
XX
AC   AAY51178;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Murine PARP1 (short) homologue protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP1.
XX
OS   Mus sp.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;  99WO-EP03889.
XX
PR   05-JUN-1998;  98DE-1025213.
PR   01-MAR-1999;  99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M,  Hoeger T,  Kroeger B,  Otterbach B,  Lubisch W,  Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44291.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
```

(Second column, rotated)

```
Qy   361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 vnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiyfasens 427

Qy   421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 ksagyvigmkcgahhvgymflgevalgrehhintdnpslksppppgfdsviarghtepdpt 487

Qy   481 QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 qdteleldgqqvvvpqggqpvpcpefssstfsqseyliyqesqcrlryllevhl 540
```

RESULT 3
```
AAU29021
ID   AAU29021 standard; Protein; 533 AA.
XX
AC   AAU29021;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human PARP-3 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001;  2001WO-US06572.
XX
PR   02-MAR-2000;  2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I,  Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45590.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 105-107; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia) , inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence  533 AA;

Query Match   99.6%;  Score 2811;  DB 22;  Length 533;
```

Left column

```
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 71-73; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   murine PARP1 protein used in the method of the invention.
XX
SQ   Sequence   528 AA;

Query Match             79.8%;  Score 2253.5;  DB 21;  Length 528;
Best Local Similarity   80.9%;  Pred. No. 1.7e-197;
Matches  432;  Conservative  37;  Mismatches  58;  Indels  7;  Gaps  4;

Qy     1 MAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60
           |||| ||  |||| ||  ||| |||||||||||||||| ||:|||||||:||  ||||
Db     1 mapkrkasvqteg--skkqrgteeedsfrstaealraapadnrvirvdpscpfsrnpgi 58

Qy    61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120
           ||||||||||:|||||||||||||||||::: :||| ||||||||||||||||||||
Db    59 qvhedydctlnqtningnnnkfyiiqlleegsrff-cwnrwgrvgevgqskmnhftcled 117

Qy   121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKV---DRGPVRVTKRV 180
           ||||||:||| | ||||  :: |||  |:||||| ||||||:||||  :---v
Db   118 akkdfkkkfwektknkweerdrfvaqpnkytlievggeaesqeavvkvdsgpvrtv---v 174

Qy   181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240
           |||:|||||||||:||||||||| | ||||:||::|||||||||||:|||||||||||:
Db   175 kpcsldpatqnlitnifskemfknamtlmnldvkkmplgkltkqqiargfealealeeam 234

Qy   241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQOPPINSPELLQAKKDMLLVLADIELAQALQA 300
           | ||||||||||| |:|||| || || ||||||||||:|||||:|||||||:||| |
Db   235 knptgdgqsleelsscfytviphnfgrsrppinspdvlqakkdmllvladielaqtlqa 294

Qy   301 V-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIW 359
           | :|||||||||||||||||||||||  :|||::||||:::||| :|:|
Db   295 apgeeeekveevphpldrdyqllrcqlqlldsgseykaiqtylkqtgnsyrcpnlrhvw 354

Qy   360 KVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 419
           |||:||| :|:|||||||:|||||||||||||||||||:|||||||||||||||||||
Db   355 kvnregedrfqahsklgnrllwhgtnvavvaailtsglrimphsggrvgkgiyfasen 414

Qy   420 SKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP 479
           ||||||:|||||||||:||| :| ||::::|| :||:||||| ||||::||||: |
Db   415 sksagyvtmhcgghqvgymflgevalgkehhitddpslksppppfdsviargqtepdp 474

Qy   480 TQDTELEDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
           ||| ||| |||:|| ||| ||| : ||||| |||||:||||||| ||
Db   475 aqdieleldqgqpvvvpgqpvvqcpsfksssfsqseyliykesqcrlrylleihl 528

RESULT 5
AAY51177
ID   AAY51177 standard; Protein; 533 AA.
XX
AC   AAY51177;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Murine PARP1 (long) homologue protein.
```

Right column

```
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP1.
XX
OS   Mus sp.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AA244290.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 67-69; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   murine PARP1 protein used in the method of the invention.
XX
SQ   Sequence   533 AA;

Query Match             79.4%;  Score 2241;  DB 21;  Length 533;
Best Local Similarity   80.1%;  Pred. No. 2.4e-196;
Matches  432;  Conservative  37;  Mismatches  58;  Indels  12;  Gaps  5;

Qy     1 MAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60
           |||| ||  |||| ||  ||| |||||||||||||||| ||:|||||||:|| : ||||
Db     1 mapkrkasvqteg--skkqrgteeedsfrstaealraapadnrvirvdpscpfsrnpgi 58

Qy    61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120
           ||||||||||:||||||||||||||||||::: :||| ||||||||||||||||||||
Db    59 qvhedydctlnqtningnnnkfyiiqlleegsrff-cwnrwgrvgevgqskmnhftcled 117

Qy   121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVK-----VDRGPVRT 175
           ||||||:||| | ||||  :: |||  |:||||| ||||||:|||||      ||  |||
Db   118 akkdfkkkfwektknkweerdrfvaqpnkytlievggeaesqeavvkalspqvdsgpvrt 177

Qy   176 VTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA 235
           : |||||||||||||:|||||||||||| | ||||:||::|||||||||||:||||||
Db   178 v---vkpcsldpatqnlitnifskemfknamtlmnldvkkmplgkltkqqiargfealea 234

Qy   236 LEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELA 295
           |||| ||||||||||||| |:|||| || || ||||||||||:|||||:||||||
Db   235 leeamknptgdgqsleelsscfytviphnfgrsrppinspdvlqakkdmllvladiela 294

Qy   296 QALQAV-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT 354
           | |||: |||||||||||||||||||||||||  :|||::
Db   296 ...
```

```
Query Match                26.1%;     Score 738;    Pred. No. 1.5e-58;
Best Local Similarity       34.3%;                   DB 22;    Length 522;
Matches 187; Conservative 103; Mismatches 191; Indels 64; Gaps 19;

Qy    16  KKKGRQA-GREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVY----EDYNCTL   70
              ||  |    |:|:  ::  :.:  |:  :|      |::|||:    ||||    ::||
Db     5  qrkgpmaggkdadrtkdnrdsvktlllkgk-apvdpecaaklgk-ahvcegddvydvml   62

Qy    71  NQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDAKKDFEKK  128
              ||||:|||||:||||||  |||||:|||:|||||||||  :|||||:|||||||:||
Db    63  nqtnlqfnnnkyylqlleddagrnfsvwmrwgrvgktgqhslvtcsgdlnkakeifqkk  122

Qy   129  FREKTKNNWAERDHFVSHPGKYTLIEV----QAEDA---QEAVVKVDRGPVRTVTKRVQ  181
              |:||||||  ||| |:|    ||||       :|:        :||   ||: |:
Db   123  fldktknnwedrenfekvpgkydmlqmdyaastqdesktkeeetlkpe----------  170

Qy   182  PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK  241
              ::     |:  ||| :   |  |:   |  |:  :|:  :|.:.:.:|:::.:|
Db   171  -sqldlrvqellklicnvqtmeemniemkydtkraplgkltvaqikagyqslkkiedcir  229

Qy   242  GPTDGGQSLELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAV  301
              || | | |:|:.::||| |  |::   |: |:| :|   :.|  : ::|:  :|
Db   230  a-gqhgralveacnefytriphdfgjsippvirtekelsdkvkllealgdiei--alklv  286

Qy   302  SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHR--CPTIQHI  358
              |:     | |  |:::| ||:||||| |:||||:|:::|||:||  |:||    ||:|||
Db   287  ksrrgle--hpldqhyrnlhcalrpldhesnefkvisqylqsthapthkdytmtlldv   343

Qy   359  WKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGIY  414
              |||||||||||||:||:|||||||||||||||||||||||||||    |||||||||||
Db   344  fevekegekeafr--edlpnrmllwhgsrlsnwvgilshglrvappeapigymfgkgiy  401

Qy   415  FASENSKSAGYVIGMKCCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH  474
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   402  fadmssksanycfasr--lkntgllllsevalgqcnelleanpkagglirgkhstkgmgk  459

Qy   475  TEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSS--------TFSQSEYLIYQESQCRLR  526
              ||||||||||||||||||||||||:||:|          |::|||:|||||:|||:|
Db   460  mapspa--------hfitlngstvplgpasdtgilnpegytlnynefivyspnqvrmr  509

Qy   527  YLLEV  531
              ||::|
Db   510  yllki  514
```

```
RESULT 7
AAB47029
ID   AAB47029 standard; Protein; 583 AA.
XX
AC   AAB47029;
XX
DT   29-MAR-2001 (first entry)
XX
DE   hPARP2.
XX
KW   Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
KW   inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
KW   infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
KW   gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
KW   endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
KW   toxic shock syndrome; multiple organ injury syndrome; vasculitis;
KW   hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
KW   eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
KW   chronic obstructive pulmonary disease; silicosis; reperfusion injury;
KW   pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
KW   bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
KW   scar tissue formation; atherosclerosis; systemic lupus erythematosus;
KW   autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
KW   graft versus host disease; allograft rejection; cystic fibrosis;
KW   chronic glomerulonephritis; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
KW   inflammatory dermatosis; contact dermatitis; atopic dermatitis;
KW   psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
```

```
Db   295  qtlqaapgeeekveevphpldrdyqllrcqlqlldsgeseykaiqtylkqtgnsyrcpn  354

Qy   355  LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIY  414
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355  lrhvwkvnqregegdrfqahsklgnrrllwhgtnvavvaailtsglrimphsggrvgkgiy  414

Qy   415  FASENSKSAGYVIGMKCCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH  474
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415  fasensksagyvtmhcgghqvgymflgevalgkehhitiddpslksppppgfdsviargq  474

Qy   475  TEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL  533
              ||||||||||||||||:||:|||||:||||||||||:|||||:||:|||||||||:||
Db   475  tepdpaqdieleldgqpvvvpqgppvqcpsfksssfsqseyliykesqcrlrylleihl  533

RESULT 6
AAU29022
ID   AAU29022 standard; Protein; 522 AA.
XX
AC   AAU29022;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Mouse PARP-2 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Mus musculus.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45597.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 109-111; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer); cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence   522 AA;
```

```
KW   Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
KW   hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
KW   leukocyte dyscrasia; thermal injury; cytokine-induced toxicity.
XX
OS   Homo sapiens.
XX
PN   WO20077179-A2.
XX
PD   21-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16629.
XX
PR   16-JUN-1999;  99US-0139543.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR   WPI; 2001-025335/03.
DR   N-PSDB; AAC85303.
XX
PT   New human poly(ADP-ribose) polymerase for treating inflammatory,
PT   neurological, cardiovascular, or neoplastic tissue growth disorders,
PT   such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
PT   metastasis
XX
PS   Claim 3; Page 94-95; 129pp; English.
XX
CC   This sequence represents human poly(ADP-ribose) polymerase (hPARP2).
CC   This protein causes the covalent addition of polymers of ADP-ribose
CC   to protein targets.  hPARP2 activity is induced in many instances of
CC   oxidative stress or during inflammation where there is direct damage
CC   to the DNA.  hPARP2 may be used to identify antagonists which
CC   may be used to treat a human having a disorder mediated by PARP2
CC   activity, such as, inflammatory, neurological, cardiovascular,
CC   or neoplastic tissue growth disorders, e.g. ischemic stroke,
CC   hemorrhagic shock, myocardial ischemia or infarction,
CC   transplantation, cerebral vasospasm; rheumatoid, osteo- or gouty
CC   arthritis, spondylitis; Behcet's disease; sepsis, septic or endotoxic
CC   shock, gram negative or positive sepsis; toxic shock syndrome; multiple
CC   organ injury syndrome secondary to septicemia, trauma, or hemorrhage;
CC   allergic or vernal conjunctivitis, uveitis, thyroid-associated
CC   ophthalmopathy; eosinophilic granuloma; asthma, chronic bronchitis,
CC   allergic rhinitis, ARDS, chronic obstructive pulmonary disease,
CC   silicosis, pulmonary sarcoidosis, pleurisy, alveolitis, vasculitis,
CC   pneumonia, bronchiectasis, pulmonary oxygen toxicity; reperfusion
CC   injury of the myocardium, brain or extremities; cystic fibrosis; keloid
CC   formation, scar tissue formation; atherosclerosis; systemic lupus
CC   erythematosus, autoimmune thyroiditis, multiple sclerosis; Reynaud's
CC   syndrome; graft versus host disease, allograft rejection; chronic
CC   glomerulonephritis; inflammatory bowel disease, Crohn's disease,
CC   ulcerative colitis, necrotizing enterocolitis; inflammatory dermatoses,
CC   contact or atopic dermatitis, psoriasis, urticaria, fever and myalgias
CC   due to infection; meningitis, encephalitis, and brain and spinal cord
CC   injury due to minor trauma; Sjogren's syndrome; diseases involving
CC   leukocyte diapedesis; alcoholic hepatitis; bacterial pneumonia;
CC   antigen-antibody complex mediated diseases; hypovolemic shock; Type 1
CC   diabetes mellitus; acute and delayed hypersensitivity; disease states
CC   due to leukocyte dyscrasia and metastasis; thermal injury; granulocyte
CC   transfusion associated syndromes; and cytokine-induced toxicity.
CC   hPARP2 and antibodies to it, can also be used to diagnose these
CC   conditions.
XX
SQ   Sequence   583 AA;
XX
```

```
Qy   59  G-TQYY----EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SK 111
             ||    :||||| |||||:|||||:|:|||||||||||:||::||:||
Db  103  gkahvycegndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhsl 162

Qy  112  INHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEAQEAVVK 167
             ||||: |||||||||: :|||||  |||| | : ||    |:::    |:||| :|||
Db  163  vacsgninkakeifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkee 222

Qy  168  VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA 227
             :|: :: : :|:| ||| | || || |   ||||
Db  223  slksplkpesq------ldlrvgeliklicnvgameemmmemkyntkkaplgkltvaqik 276

Qy  228  RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL 287
             |  | :||||    |||| | :|||
Db  277  agygslkkiedcira-gqhgralmeacnefytriphdfglrtpplirtqkelsekiqlle 335

Qy  288  VLADIELAQALQAVSEQEKTVEEVP-HPLDRDYQLLKCCLQLLDSGAPEYKVIQTYLEQT 346
             |:| || ||| |         |  ||| | |||||||||||  ||||:|| :| |:||
Db  336  algdieiaikl-----vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqst 389

Qy  347  GS---NHRCPTLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP 403
             |:    || | | || |  :|  :||  ||| |  |||  || | || |||||||||
Db  390  hapthsdytmtlldlfevekdgekeafr--edlnrmllwhgsrmsnvwgilshglriap 447

Qy  404  H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL 459
             |     |   | | |||          :::::| |||  ||||  |  |:| |
Db  448  peapitgymfgkgiyfadmsskssnycfasr--lkntglilsevalgqcnelleanpka 505

Qy  460  KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVVPQGQPVVPCPE 504
             |||||  || || | |  :  ||  || |  ::  | |  |
Db  506  egllqgkhstkglgkmappsahftvlngstvplgpasdt------qinpdgy------ 552

Qy  505  FSSSTFSQSEYLIYQESQCRLRYLLEV 531
             |::: :  || || |  | ||:|||:|
Db  553  ----tlnyneyivynpnqvrmryllkv 575
```

```
RESULT 8
AAB42909
ID   AAB42909 standard; Protein; 534 AA.
XX
AC   AAB42909;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2673 polypeptide sequence SEQ ID NO:5346.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;  99US-0127607.
PR   02-APR-1999;  99US-0127636.
```

```
Query Match             25.8%;  Score 729.5;  DB 22;  Length 583;
Best Local Similarity   35.4%;  Pred. No. 1.1e-57;
Matches  201;  Conservative  95;  Mismatches  184;  Indels  87;  Gaps  22;

Qy   15  EKKKGRQAG------REED------PFRSTA----EALKAIPAEKRIIRVDPTCPLSSNP 58
             ||   ||       ||||      |:||||    |:|| : :      :: ::
Db   46  eskkmpvaggkankdrtedkqdgmpgrswaskrvsesvkalllkgk-apvdpec--takv 102
```

```
PR   05-APR-1999; 99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shinkets RA, Leach M;
XX
DR   WPI; 2000-602362/57.
DR   N-PSDB; AAC77118.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 11; Page 4522-4524; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antiallergic; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence  534 AA;

Query Match        25.7%;   Score 726.5;  DB 21;  Length 534;
Best Local Similarity  35.1%;   Pred. No. 1.7e-57;
Matches  196;  Conservative  99;  Mismatches  187;  Indels  77;  Gaps  21;

Qy    13 GPEKKKGRQAGREED-PFRSTA----EALKAIPAEKRIIRVDPTCPLSSNPG-TQVY--  63
         ||| | | ||||| | :|  |     ||:||:|  |  :: ||: ||
Db     5 ggkankdrtedkqdgmpgrswaskrvsesvkalllkgk-apvdpec--takvgkahvyce  61

Qy    64 --EDYNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWRGVGEVGQ-SKINHFTRLE 119
           ::  |||||:||:::||::|  |:|:| : ||  | : |  :|||||::| :|
Db    62 gndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnln 121

Qy   120 DAKKDFEKKFREKTKNNWAEERDHFVSHPGKYTLIEV---QAEDEAQEAVVKVDRGPVRT 175
         || :|:||| ||||||||  :| | |||  ::::    :|
Db   122 kakeifqkfkldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkp 181

Qy   176 VTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA 235
         ::  ::||| || || |||: |:|||  : ||:| ||  ||: :: :::
Db   182 esq-----ldlrvgellklicnvqameemmemkyntkkaplgkltvaqikagyqslkk   235

Qy   236 LEEALKGPTDGGGQSLEELSSHFVTVIPHNFGHSQOPPPINSPELLQAKKDMLLVLADIELA 295
         :|: :|  |:|  ||:||||:: ||:  ::  : ||::||||::|:::||: |:|:|
Db   236 iedcira-gqhgralmeacnefytriphdfgjrtpplirtqkelsekiqllealgdieia  294

Qy   296 QALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHR 351
         || | ::| ::  |    :|||  :::|:|| | :|   ||::
Db   295 ikl-----vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdy 348

Qy   352 CPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGG 407
         ||||||   ::   |   : :| |  ||:| |||  ||   |  |||||:|   |
Db   349 tmtlldifevekdgekeafr--edlhnrmllwhgsrmsnwvgjlshgiriappeapitgy 406
```

```
Qy   408 RVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL-------- 459
         |||||||  :|||||  | | ||||  : |  :|||| : : ||
Db   407 mfgkglyfadmssksanycfasr--lkntglllsevalgqcnelleanpkaegllqgkh 464

Qy   460 -----KSPPPGFDSVIARGHTEP-DPTQDTELEDGQQVVPQGQPVPCPEFSSSTFSQ 512
              |  || ||   |||: :||  :|::  : |  || || ::  |    | | ::
Db   465 stkglgkmapssahfvtlngstvplqpasdt------gilnpdgy----------tlny 507

Qy   513 SEYLIYQESQCRLRYLLEV 531
         :|::|: :| |:|||||:|
Db   508 neyivynpnqvrmryllkv 526
```

```
RESULT          9
AAU29023
ID   AAU29023 standard; Protein; 534 AA.
XX
AC   AAU29023;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human PARP-2 protein #2.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45684.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 16; Page 125-127; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cdNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence  534 AA;

Query Match        25.7%;   Score 726.5;  DB 22;  Length 534;
Best Local Similarity  35.1%;   Pred. No. 1.7e-57;
```

RESULT 10

AAU29020

ID   AAU29020 standard; Protein; 534 AA.

XX

AC   AAU29020;

XX

DT   18-DEC-2001 (first entry)

XX

DE   Human PARP-2 protein #1.

XX

KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.

XX

OS   Homo sapiens.

XX

PN   WO200164955-A1.

XX

PD   07-SEP-2001.

XX

PF   01-MAR-2001; 2001WO-US06572.

XX

PR   02-MAR-2000; 2000US-0517467.

XX

PA   (ISIS-) ISIS PHARM INC.

XX

PI   Popoff I, Cowsert LM;

XX

DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45586.

XX

PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -

XX

PS   Example 13; Page 102-104; 168pp; English.

XX

CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.

XX

SQ   Sequence   534 AA;

Query Match   25.7%; Score 725.5; DB 22; Length 534;
Best Local Similarity   35.1%; Pred. No. 2.1e-57;
Matches 196; Conservative   99; Mismatches 187; Indels 187; Gaps   21;

Qy   13 GPEKKKGRQAGREED-PFRSTA----EALKAIPAEKRIIRVDPTCPLSSNPG-TQVY-- 63
          | |:|:|||:||||| :|:|| |:::|| ||||:| || ||:||:|
Db   5 ggkankdrtedkqdgmpgrswaskrvsesvkalllkgk-apvdpec--takvgkahvyce 61

Qy   64 --EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLE 119
            |:: |||||:|||:||||||||:::||| |: |::| ||||:::|| | |
Db   62 gndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgghslvacsgnln 121

Qy   120 DAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEAQEAVVKVDRGPVRT 175
          || |||:|||||||||||: | |||| | : || | | ::||||:|||
Db   122 kakeifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslkspikp 181

Qy   176 VTKRVQPCSLDPATQKLITNIFSKEMPKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA 235
             :|||||||| | |||:||:| ::: || | : | ||||| ||::|
Db   182 esq------ldirvgeliklicnvqameemmemkyntckkapigkltvaqikagyqslkk 235

Qy   236 LEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELA 295
          :|::: | :||| |: | ||| || ||:| :| |::| ||||:|
Db   236 iedcira-gqhgralmeacnefytriphdfgjrtpplirtqkelsekiqlleaigdieia 294

Qy   296 QALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHR 351
          :: ||||||| | || ||| :: |||:| :| : ||||:|||:||||||::| |||
Db   295 ikl-----vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdy 348

Qy   352 CPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGG 407
          |||| |||| | |: |||||||: | ||| : || | : |: |:| |
Db   349 tmtlldlfevekdgekeafr--edlhnrmllwhgsrmsnwvgilshgriahpeapitgy 406

Qy   408 RVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL------- 459
          ||||||||||:||||:|| ||||||| :|::|| || |:: |
Db   407 mfgkgiyfadmssksanycfasr--lkntglillsevalggcnelleanpkaegllqgkh 464

Qy   460 -----KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQ 512
               |||||||||||||:| ||:|| : | | : | |
Db   465 stkglgkmapssahfvtlngstvplgpasdt------gilnpdgy---------tlny 507

Qy   513 SEYLIYQESQCRLRYLLEV 531
          || ||||:||:|||||:|:|
Db   508 neyivynpnqvrmryllkv 526

```
RESULT 11
AAY51174
ID   AAY51174 standard; Protein; 570 AA.
XX
AC   AAY51174;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human brain PARP2 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M,  Hoeger T,  Kroeger B,  Otterbach B,  Lubisch W,  Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44287.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 52-54; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, the nucleic acid of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP2 protein used in the method of the invention.
XX
SQ   Sequence  570 AA;

Query Match             25.7%;  Score 725;  DB 21;  Length 570;
Best Local Similarity   35.4%;  Pred. No. 2.6e-57;
Matches  197;  Conservative  95;  Mismatches 185;  Indels  80;  Gaps  21;

Qy      15 EKKKGRQAG------REEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPG-TQVY---- 63
              | ||| ||    |  |||| ::: ||||| :  |||| ::   :   ||
Db      46 eskkmpvaggkankdrted--kqdesvkalllkgk-apvdpec--takvgkahvycegn 99

Qy      64 EDYNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDA 121
             | |||::|||||||:|||| ||:|:::|| |  |||||: :|||||:||||:|
Db     100 dvydvmlnqtnlqfnnnkyyiqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnka 159

Qy     122 KKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVKVDRGPVRTVT 177
             |:|||:|| ||||||:|||| |:|| : ||:|  :::   |:||:
Db     160 keifqkfldktknnwedrekfekvpgkydmlqmdyatntdeeetkeeslksplkpes 219

Qy     178 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQIARGFEALEALE 237
```

```
Db     220 q------idrvgeliklicnvqameemmemkyntkkaplgkitvaqikagyqslkkie 273
                    |:||   | ||||:  : || |||||       |::::::|:.|
Qy     238 EALKGPTDGGQSLELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQA 297
              : :|  |  |  :  || |||:|| ||  :|  |||| | |||||||| :|  |
Db     274 dcira-gqhgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieiaik 332
Qy     298 LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPPEYKVIQTYLEQTGS---NHRCP 353
                  |||::| |  ||::|||::  :|||   | :  :  |||:||  |:|
Db     333 l------vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytm 386
Qy     354 TLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRV 409
              ||||| | ||: |  :|  ||:|  |  |  |:  |  ||:|  |||     ||||
Db     387 tldlfevekdgekeafr--edlhnrmllwhgsrmsnwrgllshglriappeapitgymf 444
Qy     410 GKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL--------- 459
             ||||||  ||||||   |  :   ||  :: | :||| ||  :||||
Db     445 gkgiyfadmssksanycfasr--lkntglllsevalgqcnelleanpkaegllqgkhst 502
Qy     460 ---KSPPPGFDSVIARGHTEP-DPTQDTELEDGQQVVVPQGQVPCPEFSSSTFSQSE 514
                   |:  |  ||:|: | |     |:| :|| :|  : |
Db     503 kglgkmapssahfvtlngstvplgpasdt-------giinpdgy--------tlnyne 545
Qy     515 YLIYQESQCRLRVLLEV 531
             |:|| |:|  ::|||||||
Db     546 ylvynpnqvrmryllkv 562
```

```
RESULT 12
AAB11480
ID   AAB11480 standard; Protein; 570 AA.
XX
AC   AAB11480;
XX
DT   02-MAR-2001  (first entry)
XX
DE   Human brain poly-ADP-ribose-polymerase protein.
XX
KW   Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
KW   cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
KW   anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
KW   antinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
KW   2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
KW   neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
KW   Parkinson's disease; ischemic damage; microinfarction; sepsis;
KW   diabetes mellitus.
XX
OS   Homo sapiens.
XX
PN   DE19921567-A1.
XX
PD   16-NOV-2000.
XX
PF   11-MAY-1999;   99DE-1021567.
XX
PR   11-MAY-1999;   99DE-1021567.
XX
PA   (BADI ) BASF AG.
XX
PI   Lubisch W,  Sadowski J,  Kock M,  Hoeger T;
XX
DR   WPI; 2001-032983/05.
DR   N-PSDB; AAC82090.
XX
PT   Drugs for inhibiting PARP or especially homologous enzymes comprising
PT   4-substituted phthalazinone derivatives, useful e.g. for treating
PT   neurodegenerative disease, ischemic damage, tumors or diabetes -
XX
PS   Example A; Page 12-13; 14pp; German.
XX
CC   This invention describes novel 4-substituted 2H-phthalazin-1-one
CC   derivatives (I) which are used for the treatment or prophylaxis of
CC   diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;
```

CC   EC 2.4.2.30) activity. The products of the invention have nootropic,
CC   neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,
CC   cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,
CC   immunosuppressive, antinflammatory, antirheumatic, antiarthritic,
CC   antidiabetic. (I) are especially used for treating or preventing
CC   neurodegenerative disease or neuronal damage (specifically associated
CC   with ischemia, trauma or massive bleeding, especially appplexy or
CC   spinal-cranial trauma), treating or preventing ischemic damage
CC   Parkinson's disease), treating or preventing ischemic damage
CC   (specifically renal damage after renal ischemia or during and after
CC   kidney transplantation or heart damage after cardiac ischemia), treating
CC   epilepsy, specifically generalized epileptic attacks (e.g. petit mal and
CC   tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lope
CC   and complex partial attacks), treating microinfarction (e.g. during and
CC   after heart valve replacement, aneurysm resectioning and heart
CC   transplantation), revascularization of critically constricted coronary
CC   arteries (e.g. after PCTA or by-pass operations) or peripheral arteries
CC   (e.g. leg arteries), treating acute myocardial ischemia and damage during
CC   or after its mechanical or drug-induced lysis and treating tumors and
CC   their metastasis, sepsis and septic shock, inflammatory and rheumatic
CC   disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)
CC   inhibit PARP (i.e. the known form designated PARP1), they especially
CC   selectively and strongly inhibit PARP homologs, specifically the homolog
CC   PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.
CC   with K_i values of 1-20 nM) and high selectivity for PARP2 relative to
CC   PARP1 (generally by a factor of more than 5).
XX
SQ   Sequence   570 AA;

Query Match   25.7%;   Score 725;   DB 22;   Length 570;
Best Local Similarity 35.4%;   Pred. No. 2.6e-57;
Matches   197;   Conservative   95;   Mismatches   185;   Indels   80;   Gaps   21;

Qy   15 EKKKGRQAG------REEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPG-TQVY----- 63
          |:|| || |:|||| : || ::||| | | |||:||||||: ||
Db   46 eskkmpvaggkankdrted--kqdesvkalllkgk-apvdpec--takvgkahvycegn 99

Qy   64 EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDA 121
          | :| ||||||::: |||||:||||| |: ||||||||||||||::|| :
Db   100 dvydvmlngtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqqhslvacsgnlnka 159

Qy   122 KKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVKVDRGPVRTVT 177
          |:|:||| :|||||| |:|||| :|: ||| ::::   :| : |:|: ||| |
Db   160 keifqkkfldktknnwedrekfekvpgkydmlqmdyatntgdeeetkkeeslksplkpes 219

Qy   178 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLLDVKKMPLGKLSKQQIARGFEALEALE 237
          |||:| |::|| :|| : |||: | |||:||||::||||||: |::|||:||:|:
Db   220 q------ldlrvgeliklicnvqameemmmemkyntkkaplgklktvaqikagyqslkkie 273

Qy   238 EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQOPPINSPELLQAKKDMLLVLADIELAQA 297
          | :|: ||| || |||||| | ||| |||: | |: |: | ::| | : ||:| |
Db   274 dcira-gqhgralmeacnefytriphdfgjlrtpplirtqkelsekiqilealgdieialk 332

Qy   298 LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS--NHRCP 353
          |||: :||| :   :| | |:||||||| | : |:||| ::||| :| || :||||
Db   333 l-----vktelqspehplqdqhyrnlhcalrpldhesyefkvisqylqsthapthsdytm 386

Qy   354 TLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRV 409
          || ::|::||: : || |||| |||: |||| ::| |||::|||:|| : |||||
Db   387 tlldlfevekdgekeafr--edlnrmllwhgsrmsnwvgilshglrlappeapitgymf 444

Qy   410 GKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL---------- 459
          |||||| :||||| : |: :| ||| | :| || :| :
Db   445 gkgiyfadmnssksanycfasr--lkntgllllsevalggcnelleanpkaegllqgkhst 502

Qy   460 ---KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSE 514
          : | | ::| | | || ::| | |:   :|
Db   503 kglgkmapssahfvlngstvplgpasdt------gilnpdgy-------tlnyne 545

Qy   515 YLIYQESQCRLRYLLEV 531
          |||||| ||||:|||||
Db   546 ylvynpnqvrmryllkv 562

RESULT 13
AAB60693
ID   AAB60693 standard; Protein; 521 AA.
XX
AC   AAB60693;
XX
DT   11-MAY-2001 (first entry)
XX
DE   Human poly(ADP-ribose) synthetase sbhPARS2.
XX
KW   Human; poly(ADP-ribose) synthetase; sbhPARS2; vaccine; drug screening;
KW   ischaemic disorder; cerebral ischaemia; cardiac ischaemia;
KW   myocardial infarction; stroke; inflammation; autoimmune disease;
KW   diabetes; multiple sclerosis; neurodegenerative disease;
KW   Parkinson's disease; Alzheimer's disease; chromosome localisation.
XX
OS   Homo sapiens.
XX
PN   WO200112645-A1.
XX
PD   22-FEB-2001.
XX
PF   10-AUG-2000; 2000WO-US21775.
XX
PR   12-AUG-1999;   99US-0373441.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI   Barone F, Field J, Kabnick KS, Li X, McQueney MS, Zhu Y;
XX
DR   WPI; 2001-211196/21.
DR   N-PSDB; AAF59996.
XX
PT   New human polypeptide of the polyADPribose synthetase family for use in
PT   diagnostic assays and for screening modulators used for preventing and
PT   treating inflammation, autoimmune disease and Alzheimers disease -
XX
PS   Claim 1; Page 21; 30pp; English.
XX
CC   This sequence represents a novel human poly(ADP-ribose)
CC   synthetase, sbhPARS2. The invention also relates to fragments,
CC   variants and sequences with at least 95% identity to the sbhPARS2
CC   protein or nucleotide sequence; expression systems and host cells
CC   comprising an sbhPARS2 nucleic acid sequence; the recombinant expression
CC   of sbhPARS2; and an antibody specific for sbhPARS2. sbhPARS2 proteins
CC   and nucleotides are useful as vaccines for inducing an immunological
CC   response in a mammal. The sbhPARS2 protein is useful for identifying
CC   compounds which inhibit or stimulate its activity or expression level.
CC   Such agonists and antagonists of sbhPARS2 are useful for treating human
CC   diseases including ischaemia and ischaemic tissue injury (e.g., cerebral
CC   and cardiac ischaemia, myocardial infarction, stroke), inflammation,
CC   autoimmune disease (e.g. diabetes, multiple sclerosis) and
CC   neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's
CC   disease). sbhPARS2 nucleic acids are useful as diagnostic reagents for
CC   detecting mutations in the associated gene; as hybridisation probes
CC   to isolate full-length sbhPARS2 cDNAs and sbhPARS2 genomic clones; and
CC   for chromosome localisation studies. The sbhPARS2 protein is also useful
CC   as an immunogen to produce antibodies for therapeutic use. sbhPARS2
CC   proteins, nucleotides and antibodies are also useful in screening methods
CC   for detecting the effect of added compounds on the production of mRNA and
CC   protein in cells.
XX
SQ   Sequence   521 AA;

Query Match   25.6%;   Score 722;   DB 22;   Length 521;
Best Local Similarity 34.7%;   Pred. No. 4.3e-57;
Matches   192;   Conservative   99;   Mismatches   184;   Indels   78;   Gaps   20;

Qy   13 GPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPG-TQVY----EDYN 67
          |: | || ::: |::|||:|| ||: ||||||| : ||| | :|:

```
DR   N-PSDB; AAZ60617.
XX
PT   Modulating cell death, growth and stress resistance in eukaryotes,
PT   specifically plants, used, e.g. to impart fungus or nematode resistance
PT   -
XX
PS   Claim 18; Page 99-101; 126pp; English.
XX
CC   The present sequence represents a NAP protein. This protein is a
CC   poly(ADP-ribose) polymerase (PARP) protein (also known as
CC   poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
CC   cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide
CC   sequences can be used for modulation of programmed cell death in
CC   eukaryotic cells. The method is used, specifically in plants, to induce,
CC   or protect against, programmed cell death, depending on the extent to
CC   which PARP activity is reduced. Reducing expression of endogenous NAP
CC   class PARP only is also used to modulate programmed cell death, to
CC   increase growth rate and to produce plant cells that are more tolerant
CC   of stress (cold, chemical treatments, pathogens, pests, drought, heat,
CC   etc., or during transformation). Particular applications are generation
CC   of plants that are resistant to fungi or nematodes; are male or female
CC   sterile; or have better seed-shatter properties. The methods are also
CC   used to improve growth of transformed plant cells (and derived calli or
CC   complete plants).
XX
SQ   Sequence  637 AA;

Query Match             24.4%;  Score 688.5;  DB 21;  Length 637;
Best Local Similarity   36.2%;  Pred. No. 7e-54;
Matches  174;  Conservative  89;  Mismatches 185;  Indels  33;  Gaps  16;

Qy       66 YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK 123
            |:  |||||| :||||||||||  :|||||||||||||||||| ||||: ||||  : :|
Db      169 ydailnqtnvrdnnnkffivlqvlesdskktymvytrwgrvgvkggskldgpydswdraie 228

Qy      124 DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQP- 182
            |  |||||||:|||  |||||:||||:||||:|: ||||| :|:   |||:||||:: :|
Db      229 iftnkfndktknywsdrkefiphpksytwlemdygkeendspvnnd---ipsssevkpe 285

Qy      183 -CSLDPATQKLITNIFSKEMPFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 241
             ||| :||||||| :||||| |||| ||||:||||  :|| ||  :::|  ||:|:|:||
Db      286 gskldtrvakfisllcnvsnmaqhnmeigynanklplgkiskstiskgyevlkrisevi- 344

Qy      242 GPTDGGQSLEELSSHFYTVIPHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ 299
            |||||||||||||||||:||||| ||||||   ||||||||||||:|||||||||||||:
Db      345 -drydrtrleelsgefytviphdfgfkkmsgfvidtpqklkqkienvealgeielatkll 403

Qy      300 AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNRHRCPTLQ- 356
            || |||||:|  |  :|  |     :||  :||||||||||||||||  |:|  ||||:|
Db      404 svdpglq-----ddplyyhyqqlnclgltpvgndseefsmvanymenthakthsgytveia 458

Qy      357 HIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKG 412
            || ||||:|  :|| ||:|||:||| || :::|  |::|||| ::||     |||:|||
Db      459 qifrasraveadrfqqfsssknrmllwhgsritnwagilsqglriappeapvtgymfgkg 518

Qy      413 IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 472
            ||| ||||:||||||||||||:|||||||| ||:|:||:|||||||||:||| || :||
Db      519 vyfadmfsksanycya-ntgand-gvlllcevalgdmnellysdynadnlppgklstkgv 576

Qy      473 GHTEPDPTQDTELELDGQQVVPQGQPV--PCPEFSSSTFSQSEYLIYQESQCRLRYLLE 530
            | | :|:||:: | || :|:    |    |:|||:||||||||| ::|| : :|||||
Db      577 gktapnpseaqtle-dg--vvvplgkpversc--skgmllyneyivynveqikmryviq 630

Qy      519 QESQCRLRYLLEV 531
            | |:|||||||:|
Db      501 npnqvrmrylikv 513

Qy      531 v 531
            |
Db      631 v 631
```

```
RESULT 14
AAY68835
ID   AAY68835 standard; Protein; 637 AA.
XX
AC   AAY68835;
XX
DT   16-MAY-2000  (first entry)
XX
DE   The poly(ADP-ribose) polymerase NAP protein of Arabidopsis.
XX
KW   NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
KW   programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
KW   pest; drought; heat; fungi; nematode; seed-shatter.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..138
FT                   /note= "these residues are specifically claimed in
FT                   claim 18"
XX
PN   WO200004173-A1.
XX
PD   27-JAN-2000.
XX
PF   12-JUL-1999;  99WO-EP04940.
XX
PR   17-JUL-1998;  98US-0118276.
XX
PA   (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX
PI   Babychuk E, Kushnir S, De Block M;
XX
DR   WPI; 2000-182436/16.
```

Left column (continued alignment / sequence):

```
Db        5 ggkankdrtedkqd------esvkalllkgk-apvdpec--takvgkahvycegndvyd 54

Qy       68 CTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDAKKDF 125
            |||||||:|||  |||||||| :||||||||||||||||||| |||:  ||||  : :|
Db       55 vmlnqtnlqfnnnkyyliqlleddaqrnfsvwnrwgrvgkmgghslvacsgnlnkakeif 114

Qy      126 EKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVKVDRGPVRTVTKRVQ 181
            :||||||||| ::|| || |: :  |||||     :|  :|  ::::  |:: :: :: :
Db      115 qkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkeeslkspikpesq--- 171

Qy      182 PCSLDPATQKLITNIFSKEMPFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEEEALK 241
            | ||: :|||||| |||| ||| ||||| ||||::|:: ||:|:|:||||:||:| ::
Db      172 --ldlrvgeliklicnvqameenmmenkyntkkaplgkltvaqikagyqslikkiedcir 228

Qy      242 GPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAV 301
            |||||||||||||| || |||||||| : || |||||||| ||:||:|||||||||||:|
Db      229 a-ghgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieiaikl--- 284

Qy      302 SEQKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS--NHRCPTLQH 357
            ||:|:|||  || :|| ::: ||:: :|| :|||||||||:: || :   |:||||::|
Db      285 ---vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytmtlld 341

Qy      358 IWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGI 413
            || |||:|  :|| ||:|||:||| || :::|  |::|||| ::||     |||:|||
Db      342 lfevekdgekeafr--edlhnrmllwhgsrmsnwvgilshglriiappeapitgymfgkgi 399

Qy      414 YFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL-----459
            ||| ||||:||:|| :|||||:|||||:|:||:|:||:|||||||||
Db      400 yfadmssksanycfasr--lkntglllsevalggcnelleanpkaegllqgkhstkglg 457

Qy      460 KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYLIY 518
            |  ||  : : | : |  :  |:| ||| ||       |  :|  : |||||||||:|
Db      458 kmapssahfvtlngstvplgpasdt------gilnpdy--------tlnyneyivy 500

Qy      519 QESQCRLRYLLEV 531
            | |:|||||||:|
Db      501 npnqvrmrylikv 513
```

```
RESULT 15
AAB93513
ID   AAB93513 standard; Protein; 531 AA.
XX
```

```
AC    AAB93513;
XX
DT    26-JUN-2001  (first entry)
XX
DE    Human protein sequence SEQ ID NO:12844.
XX
KW    Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS    Homo sapiens.
XX
PN    EP1074617-A2.
XX
PD    07-FEB-2001.
XX
PF    28-JUL-2000; 2000EP-0116126.
XX
PR    29-JUL-1999;    99JP-0248036.
PR    27-AUG-1999;    99JP-0300253.
PR    11-JAN-2000;    2000JP-0118776.
PR    02-MAY-2000;    2000JP-0183767.
PR    09-JUN-2000;    2000JP-0241899.
XX
PA    (HELI-) HELIX RES INST.
XX
PI    Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI    Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR    WPI; 2001-318749/34.
XX
PT    Primer sets for synthesizing polynucleotides, particularly the 5602
PT    full-length cDNAs defined in the specification, and for the detection
PT    and/or diagnosis of the abnormality of the proteins encoded by the
PT    full-length cDNAs -
XX
PS    Claim 8; SEQ ID 12844; 2537pp + CD ROM; English.
XX
CC    The present invention describes primer sets for synthesising 5602
CC    full-length cDNAs defined in the specification, where a primer set
CC    comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC    to the complementary strand of a polynucleotide which comprises one of
CC    the 5602 nucleotide sequences defined in the specification, where the
CC    oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC    of an oligonucleotide comprising a sequence complementary to the
CC    complementary strand of a polynucleotide which comprises a 5'-end
CC    sequence and an oligonucleotide comprising a sequence complementary to a
CC    polynucleotide which comprises a 3'-end sequence, where the
CC    oligonucleotide comprises at least 15 nucleotides and the combination of
CC    the 5'-end sequence/3'-end sequence is selected from those defined in
CC    the specification. The primer sets can be used in antisense therapy and
CC    in gene therapy. The primers are useful for synthesising polynucleotides,
CC    particularly full-length cDNAs. The primers are also useful for the
CC    detection and/or diagnosis of the abnormality of the proteins encoded by
CC    the full-length cDNAs. The primers allow obtaining of the full-length
CC    cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC    AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC    AAH13632 represent human amino acid sequences; and AAH13629 to AAH13632
CC    AAB95893 represent oligonucleotides, all of which are used in the exemplification
CC    of the present invention.
XX
SQ    Sequence   531 AA;
XX

Query Match        24.2%;    Score 682.5;  DB 22;  Length 531;
Best Local Similarity  36.9%;    Pred. No. 1.9e-53;
Matches  184;  Conservative  86;  Mismatches 173;  Indels  55;  Gaps  18;

Qy      15 EKKKGRQAG-----REED-----PFRSTA-----EALKAIPAEKRIIRVDPTCPLSSNP 58
               || |||   |    ||     |:||::    :|  || :|| :||: :
Db      46 eskkmpvaggkankdrtedkqdgmpgrswaskrvsesvkalllkgk-apvdpec--takv 102

Qy      59 G-TQVY----EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SK 111
            | |||     || | ||:|| |||||||::|||||: | ||||||||:||:::|
Db     103 gkahvycegndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvmwrwgrvgkmngqhsl 162
```

```
Qy     112 INHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVK 167
            : ||:|   ||| ||::|| :|||||| :|||| :  ||| :   ||  :
Db     163 vacsgnlnkakeifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkee 222

Qy     168 VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA 227
            : |: :   :  :          |: ||:|         :| |||| |
Db     223 slksplkpesq------ldlrvqeliklicnvgameemmemkyntkkaplgkltvaqik 276

Qy     228 RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL 287
            ||:||::||: ::|: | ||||| ||:|||:  :|  :
Db     277 agyqslkkiedcira-gqhgralmeacnefytriphdfglrtpplirtqkelsekiqlle 335

Qy     288 VLADIELAQALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 346
            | |||:|   |  || ||: |  : ||||||:||||||::|||||||| :
Db     336 algdielaikl-----vktelqspehpldqhyrnlhcalrpidhesyefkvisqylqst 389

Qy     347 GS---NHRCPTLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP 403
            :     :|: :::||::| :||:|  |: || | :  |::|  :||||:|||||||
Db     390 hapthsdytmtlldifevekdgekeafr--edlnrmllwhgsrmsnwvgilshglriap 447

Qy     404 H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL 459
            :  :|  ||||||| :|||||  :|||||  ||||  :  :::  ||
Db     448 peapitgymfgkgiyfadmssksanycfasr--lkhtglll1sevalgqcnelleanpka 505

Qy     460 KSPPPGFDSVIARGHTEP 477
            |    : |  :   | ||
Db     506 egllqgkhstkglgkmap 523
```

Search completed: August 29, 2002, 07:57:14
Job time: 291 sec

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: August 29, 2002, 07:58:08 ; Search time 46.79 Seconds
                                            (without alignments)
                                            278.240 Million cell updates/sec

Title: US-09-701-586B-4
Perfect score: 2823
Sequence: 1 MAPKPKPWVQTEGPEKKKGR.........EYLIYQESQCRLRYLLEVHL 533

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
               1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
               2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
               3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
               4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
               5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
               6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	668.5	23.7	1014	4	US-09-078-347A-3	Sequence 3, Appli
2	666.5	23.6	1013	4	US-08-860-886-2	Sequence 2, Appli
3	138.5	4.9	1327	4	US-09-196-387-2	Sequence 2, Appli
4	113.5	4.0	1164	4	US-08-923-992A-2	Sequence 2, Appli
5	112.5	4.0	717	4	US-08-910-925-1	Sequence 1, Appli
6	112.5	4.0	1098	4	US-08-923-992A-8	Sequence 8, Appli
7	110.5	3.9	1128	4	US-08-923-992A-6	Sequence 6, Appli
8	109	3.9	743	4	US-08-910-925-3	Sequence 3, Appli
9	108.5	3.8	1871	2	US-08-694-869-1	Sequence 1, Appli
10	108.5	3.8	1871	3	US-09-349-546-1	Sequence 1, Appli
11	108.5	3.8	2548	4	US-09-172-422-1	Sequence 1, Appli
12	107.5	3.8	1164	4	US-08-923-992A-10	Sequence 10, Appl
13	105.5	3.7	1464	4	US-09-045-360-2	Sequence 2, Appli
14	105	3.7	641	1	US-08-961-083-160	Sequence 160, App
15	103.5	3.7	1528	1	US-08-326-117B-2	Sequence 2, Appli
16	103.5	3.7	1528	1	US-08-982-129-2	Sequence 2, Appli
17	102.5	3.6	2482	1	US-08-328-254-6	Sequence 6, Appli
18	101.5	3.6	1104	2	US-08-923-992A-4	Sequence 4, Appli
19	101.5	3.6	1461	2	US-08-993-228-10	Sequence 10, Appl
20	101.5	3.6	3075	2	US-08-460-309-5	Sequence 5, Appli
21	101.5	3.6	3075	2	US-08-125-077-5	Sequence 5, Appli
22	101	3.6	452	2	US-08-686-599A-18	Sequence 18, Appl
23	101	3.6	493	2	US-08-686-599A-5	Sequence 5, Appli
24	101	3.6	493	2	US-08-686-599A-16	Sequence 16, Appl
25	100	3.5	443	2	US-08-795-475-6	Sequence 6, Appli
26	100	3.5	467	2	US-08-686-599A-17	Sequence 17, Appl
27	100	3.5	737	1	US-08-185-432-2	Sequence 2, Appli
28		3.5	100	737	1 US-08-185-432-4	Sequence 4, Appli
29		3.5	100	1786	4 US-08-973-462-8	Sequence 8, Appli
30		3.5	99.5	3248	1 US-08-353-700-1	Sequence 1, Appli
31		3.5	99.5	3248	5 PCT-US95-16216-1	Sequence 1, Appli
32		3.5	98	1939	4 US-09-310-187A-1	Sequence 1, Appli
33		3.4	97.5	163	2 US-08-044-618-6	Sequence 6, Appli
34		3.4	96	2391	2 US-08-446-855A-2	Sequence 2, Appli
35		3.4	96	2391	4 US-09-150-741-2	Sequence 2, Appli
36		3.4	95	472	1 US-08-021-601-8	Sequence 8, Appli
37		3.4	95	472	1 US-08-082-849B-8	Sequence 8, Appli
38		3.4	95	472	5 PCT-US94-01624-8	Sequence 8, Appli
39		3.3	94.5	456	1 US-08-021-601-6	Sequence 6, Appli
40		3.3	94.5	456	1 US-08-082-849B-6	Sequence 6, Appli
41		3.3	94.5	456	5 PCT-US94-01624-6	Sequence 6, Appli
42		3.3	94	1066	4 US-09-541-782-8	Sequence 8, Appli
43		3.3	93	504	3 US-09-231-529-6	Sequence 6, Appli
44		3.3	93	504	4 US-08-977-816-6	Sequence 6, Appli
45		3.3	92.5	2285	4 US-09-308-375-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-078-347A-3
; Sequence 3, Application US/09078347A
; Patent No. 6132968
; GENERAL INFORMATION:
; APPLICANT: Le, Xiao-Chun
; APPLICANT: Weinfeld, Michael
; APPLICANT: Xing, James Z.
; TITLE OF INVENTION: Methods for Quantitating Low Level
; TITLE OF INVENTION: Modifications of Nucleotide Sequences
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Medlen & Carroll, LLP
;   STREET: 220 Montgomery Street, Suite 2200
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94104
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/078,347A
;   FILING DATE: 13-MAY-1998
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Carroll, Peter G.
;   REGISTRATION NUMBER: 32,837
;   REFERENCE/DOCKET NUMBER: UALB-03283
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 705-8410
;   TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1014 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: not relevant
;   TOPOLOGY: unknown
;   MOLECULE TYPE: protein

US-09-078-347A-3

Query Match   23.7%; Score 668.5; DB 4; Length 1014;
Best Local Similarity   33.4%; Pred. No. 2.4e-59;
Matches 187; Conservative 100; Mismatches 204; Indels 69; Gaps 24;

Qy   7 PW-------VQTEGPEKKKGRQAGR------EEDPFRSTAEALKAIPAEKRIIRVDPTC 52
          || |:: | || || | | : || : :|| | |||

```
Db   480  PWGAEVKAEPVEVVAPRGKSGAALSKKSKGQVKEEGINKSEKRMKL--TLKGGAAVDPDS  537

Qy    53  PLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWGRVGE   106
               |  :    : |  ||       :  ||  | :    ||   ||||
Db   538  GLEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYW--IFRSWGRVGT   594

Qy   107  V-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEA   164
          |  :  :      ||| : |  | : ||  |  :   :||  ||  :|   ||
Db   595  VIGSNKLEQMPSKEDAIEQFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EA   651

Qy   165  VVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQ   224
          | :|:   |  ||  ||    : | |  ||| ::| || || :|:||||||||||:|
Db   652  VKKLTVNP-GTKSKLPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKR   704

Qy   225  QIARGFEALEAEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKD   284
          ||| | :  :|:  ||      :||:|| :||| | |||:|   ||  :|| :|||
Db   705  QIQAAYSILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVE   763

Qy   285  MLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLE   344
           | | | |||   |     | |  :   ::  |: ||  ::||  | :   ||:|
Db   764  MLDNLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEIIRKYVK   819

Qy   345  QT-GSNHRCPTLQ--HIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRI   401
          || |||||| |:  |||| |  :||  |:| :: : |||:||| : |   ||||:|||
Db   820  NTHATTHNAYDLEVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRI   879

Qy   402  MPH----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH-----VGYMFLGEVALGREHH   451
          | |      |: ||||||||  | |||   |:|      ||||||  ||:|
Db   880  APPEAPVTGYMFGKGIYFADMVSKSANY----CHTSQGDPIGLILLGEVALGNMYELKH   931

Qy   452  INTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSSTFS   511
          |    | || |   ||   | :  | |  || |||: ::  | |   | :
Db   932  LKHAS-HISRLPKGHSVKGLGKTTPDPS--ANISLDG--VDVPLGTGISSGVIDTSLL-   985

Qy   512  QSEYLIYQESQCRLRYLLEV   531
          ::|| ||  | |||:|||:
Db   986  YNEYIVYDIAQVNLKYLLKL  1005
```

RESULT 2
US-08-860-886-2
; Sequence 2, Application US/08860886
; Patent No. 635009
; GENERAL INFORMATION:
;   APPLICANT: Burkle, Alexander
;   APPLICANT: zur Hausen, Harald
;   APPLICANT: Jan-Heiner, Kupper
;   TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE
;   TITLE OF INVENTION: IN GENE THERAPY
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Pennie & Edmonds, LLP
;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 10036-2811
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: Windows
;   SOFTWARE: FastSEQ for Windows Version 2.0b
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/860,886
;   FILING DATE: 03-OCT-1997
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Coruzzi, Laura A
;   REGISTRATION NUMBER: 30,742
;   REFERENCE/DOCKET NUMBER: 8484-0028-999

;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650-493-4935
;   TELEFAX: 650-493-5556
;   TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1013 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   FRAGMENT TYPE: internal
US-08-860-886-2

Query Match   23.6%; Score 666.5; DB 4; Length 1013;
Best Local Similarity   33.4%; Pred. No. 3.9e-59;
Matches 186; Conservative 101; Mismatches 207; Indels 63; Gaps 24;

```
Qy     7  PW--------VQTEGEKKKGRQAGR-----EEDPRSTAEALKAIPAEKRIIRVDPTC   52
          ||         :         ||            |:               :  ||
Db   479  PWGAEVKAEPVEVVAPRGKSGAALSKKSKGQVKEEGINKSEKRMKL--TLKGGAAVDPDS  536

Qy    53  PLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWGRVGE   106
               |  :    : |  ||       :  ||  | :    ||   ||||
Db   537  GLEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYW--IFRSWGRVGT   593

Qy   107  V-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEA   164
          |  :  :      ||| : |  | : ||  |  :   :||  ||  :|   ||
Db   594  VIGSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EA   650

Qy   165  VVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQ   224
          | :|:   |  ||  ||    : | |  ||| ::| || || :|:||||||||||:|
Db   651  VKKLTVNP-GTKSKLPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKR   703

Qy   225  QIARGFEALEAEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKD   284
          ||| | :  :|:  ||      :||:|| :||| | |||:|   ||  :|| :|||
Db   704  QIQAAYSILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVE   762

Qy   285  MLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLE   344
           | | | |||   |     | |  :   ::  |: ||  ::||  | :   ||:|
Db   763  MLDNLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEIIRKYVK   818

Qy   345  QT-GSNHRCPTLQ--HIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRI   401
          || |||||| |:  |||| |  :||  |:| :: : |||:||| : |   ||||:|||
Db   819  NTHATTHNAYDLEVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRI   878

Qy   402  MPH----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH-----VGYMFLGEVALGREHHINT   454
          | |      |: ||||||||  | |||   |:|      ||||||  ||:|
Db   879  APPEAPVTGYMFGKGIYFADMVSKSANY----CHTSQGDPIGLILLGEVALGNMYELKH   933

Qy   455  DNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSSTFSQSE   514
          |    | || |   ||   | :  | |  || |||: ::  | |   | :
Db   934  AS-HISKLPKGHSVKGLGKTTPDPS--ANISLDG--VDVPLGTGI-SSGVNDTSLLYNE   987

Qy   515  YLIYQESQCRLRYLLEV   531
          ::|| ||  | |||:||:
Db   988  YIVYDIAQVNLKYLLKL  1004
```

RESULT 3
US-09-196-387-2
; Sequence 2, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
;   APPLICANT: de Lange, Titia
;   APPLICANT: Smith, Susan
;   TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
;   TITLE OF INVENTION: OF USE THEREOF
;   NUMBER OF SEQUENCES: 12
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Klauber & Jackson
;   STREET: 411 Hackensack Avenue, 4th Floor

```
  CITY: Hackensack
  STATE: New Jersey
  COUNTRY: USA
  ZIP: 07601
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/196,387
  FILING DATE:
  CLASSIFICATION:
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 09/095,225
  FILING DATE: June 10, 1998
 ATTORNEY/AGENT INFORMATION:
  NAME: Jackson Esq., David A.
  REGISTRATION NUMBER: 26,742
  REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: 201-487-5800
  TELEFAX: 201-343-1684
  TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
  LENGTH: 1327 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
US-09-196-387-2

Query Match      4.9%;     Score 138.5;    DB 4;    Length 1327;
Best Local Similarity 24.4%;    Pred. No. 0.00011;
Matches  73;  Conservative  41;  Mismatches 102;  Indels 83;  Gaps 16;

Qy  277  ELLQAKKDMLLVLADI--ELAQ-ALQAVSEQEKTVEEVPHPL-----DRDYQLLKCQLQ  327
Db 1045  DIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQ 1104

Qy  328  ---LLDSGAP---EYKVIQTYLEQTGSNHRCP---------TLQHIWKVNQEGEEDRF-  370
Db 1105  GTILLDL-APEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFC 1163

Qy  371  ----QAHSKLGNRKLLWHGTNMAVVAAILTSGL-RIMPHSGGRVGKGIYFASENSKS  422
Db 1164  HRQKEVSEENHNHHNERMLFHGS--PFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKS 1221

Qy  423  AGYVIGM-----KCGAH------HVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVI  470
Db 1222  NQYVIGIGGTGCPTHKDRSCYICHRQMLFC-RVTLGKS-FLQFSTMKMAHAPPGHHSVI 1279

Qy  471  ARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLL  529
Db 1280  ---------GRP------SVNGLAYAEVVIYRGEQAYPEYLI 1306


RESULT  4
US-08-923-992A-2
; Sequence 2, Application US/08923992A
; Patent NO. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
```

```
  CITY: Washington
  STATE: D.C.
  COUNTRY: USA
  ZIP: 20005
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/923,992A
  FILING DATE: 05-SEP-1997
  CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 60/024,707
  FILING DATE: 06-SEP-1996
 ATTORNEY/AGENT INFORMATION:
  NAME: Esmond, Robert W.
  REGISTRATION NUMBER: 32,893
  REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 371-2600
  TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
  LENGTH: 1164 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
 MOLECULE TYPE: protein
US-08-923-992A-2

Query Match      4.0%;     Score 113.5;    DB 4;    Length 1164;
Best Local Similarity 20.2%;    Pred. No. 0.034;
Matches 107;  Conservative  72;  Mismatches 205;  Indels 147;  Gaps 25;

Qy   21  QAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN  79
Db  309  QATQVKNQFLENAQKLKEIQ-----PLIKETNVKLYKAMSESSLEQVEKELKHN 356

Qy   80  NKFYIIOLLQDSNRFFTCW-----NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREK--- 132
Db  357  SEANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFKTSEQV 416

Qy  133  ------------TKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPV--RTVT 177
Db  417  TPKKRVKRDLAANENNQQKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDFSDLL 473

Qy  178  KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 237
Db  474  TKYNPSVSDRISTNYKTN-----TDNHKIAEITIKNL--KLNESQTV- 513

Qy  238  EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQA 297
Db  514  -TLKAKDDSGNVVEKT----FTITVQKKEEKQVP--KTPE----QKD- 549

Qy  298  LQAVSEQEKTVEVP-HPLDRDYQLLKCQLQ-LLDSGAPEYKVIQTYLEQTGSNHRCPTL 355
Db  550  -------SKTEEKVPQEPKSND---KNQLQELIKSAQQELEKLEKAIKEL-- 589

Qy  356  QHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYF 415
Db  590  ------MEQPEIPSNPEYGIQKSIWESQKEPIQEAI-TSFKKIIGDSSSKYYTEHYF 639

Qy  416  ASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHIN--TDNPSLKSPPPGFDSVIAR- 472
Db  640  NKYKSDFMNYQL-----HAQMEMLTRKVV----QYMNKYPDNAEIKKI--FESDMKRT 686

Qy  473  ------GHTEPD-------PTQDTELELDGQQVVPQGQPVPCPEFS 506
Db  687  KEDNYGSLENDALKGYFEKYFLTPFNKIKQIVDDLDKKVEQDQPAPIPENS 737
```

RESULT 5
US-08-910-925-1
; Sequence 1, Application US/08910925
; Patent No. 6162601
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,925
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0365 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 717 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: FIBRNOT01
; CLONE: 53219
US-08-910-925-1

Query Match   4.0%; Score 112.5; DB 4; Length 717;
Best Local Similarity 19.3%; Pred. No. 0.019;
Matches 102; Conservative 81; Mismatches 229; Indels 117; Gaps 21;

Qy   13 GPEKKKGRQAGREEDP---FRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNC 68
         | |:::|:::.:| || : | :: |:|
Db   87 GGERRTRRESRQESDPEDDDVKKPALQSSVVATSKERTRRD----------LIQDQN- 133

Qy   69 TLNQTNIENNNKFYII-----QLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKK 123
         ::: .:| ||| : | |::: :|
Db 134 -MDEKGKQRNRRIFGLLMGTLQKFKQEST---VATERQKRRQEIEQ-KLE--VQAEEERK 186

Qy 124 DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAE-DEAQEAVVKVDRGPVRTVTKR--- 179
         | :: | |: || | | ::: |
Db 187 QVENERRELFEERRAKQTELRLLEQKVELAQLQFEWNEHNAKIIKY----IRTKTKPHLF 242

Qy 180 VQPCSLDPATQKLITNIFSK--EMFKNTMALMDLDVKKMPLG------KLSKQQIARGFE 231
         |||||||| :|| | :: :||
Db 243 YIPGRMCPATQKLIEESQRKMNALFEGRRIEFAEQINKMEARPRRQSMKEKEHQVVRNEE 302

Qy 232 ALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLAD 291
         || : ||| ::: | ||| : ::|
Db 303 QKAEQEEGKVAQRE--EELEETGNQHNDVEIEFAGEEE----------EKEIAIVHSD 348

Qy 292 IELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHR 351
         | ::| .:| :|| ::| |
Db 349 AEKEQ-----EEEQKQEMEVKMEEETEVRESEKQ-QDSQPEEVMDVLEMVE------- 394

Qy 352 CPTLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGK 411
         : :|:
Db 395 --NVKHVIADQEVMETNRVESVEPSENE---------------ASKELEPE------- 428

Qy 412 GIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIA 471
         :
Db 429 -MEFEIEPDK--------ECKSLSPGKENVSALDMEKESEEKEKESEPQPEP------- 472

Qy 472 RGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLLYQE 520
         : | ::: | ::| || ||
Db 473 --VAQPQSQPQLQLQSQSQPVLQSQPSQPEDLSLAVLQPTPQVTQE 519

RESULT 6
US-08-923-992A-8
; Sequence 8, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

Query Match   4.0%; Score 112.5; DB 4; Length 1098;
Best Local Similarity 20.0%; Pred. No. 0.039;
Matches 107; Conservative 75; Mismatches 199; Indels 153; Gaps 26;

Qy   21 QAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN 79
         || :: :| ::| || :||| |:
Db 272 QATQVKNQFLENAQKLKEIQ----------PLIKETNVKLYKAMSESLEQVEKELKHN 319

Qy   80 NKFYIIQLLQDSNRFFTCW----NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREK--- 132
         | | :: | | :| : ::|:
Db 320 SEANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFKTSEQV 379

```
Query Match                3.9%;    Score 110.5;   DB 4;    Length 1128;
Best Local Similarity     20.0%;    Pred. No. 0.065;
Matches 106; Conservative 73; Mismatches 205; Indels 147; Gaps 25;

Qy     21 QAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN 79
             ||:||:|::| |:||     |        ::|: ::|| | :| :
Db    273 QATQVKNQFLENAQKLKEMQ-------------PLIKETNVKLYKAMSESLEQVEKELKHN 320

Qy     80 NKFYIIQLLQDSNRFFTCW----NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREK---- 132
             :: :|||||| ||: |:|    || :  | :::|  || |::| ||:|| ::| |||:
Db    321 SEANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFKTSEQV 380

Qy    133 --------TKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPV---RTVT 177
                      :|| :| || :| ||:|||| ||:| |  |:|||    ||:|
Db    381 TPKKRVKRDLAANENNQQKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDFSDLL 437

Qy    178 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 237
             ||     |   |   ||   :|     ||:| ::| :: ||:| :
Db    438 TKYNPSVSDRISTNYKTN------TDNHKIAEITIKNL---KLNESQTV---------- 477

Qy    238 EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQA 297
             |:|   |   ||  :|  || |   ||  || |:|  |||   | |   :|:
Db    478 -TLKAKDDSGNVVEKT----FTITVQKKEEKQVP--KTPE----QKD----------- 513

Qy    298 LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQ-LLDSGAPEYKVIQTYLEQTGSNHRCPTL 355
             | | ||:  |  :|  ||| || || |:|   |  :|| :|| |::|:|: |  |::|
Db    514 ------SKTEEKVPQEPKSND----KNQLQELIKSAQQELEKLEKAIKEL-------- 553

Qy    356 QHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYF 415
             ::  :|  :|: :: :|||  |: | :  ||   || : | ::
Db    554 -------MEQPEIPSNPEYGIQKSIWESQKEPIQEAI-TSFKKIIGDSSSKYYTEHYF 603

Qy    416 ASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHIN-TDNPSLKSPPPGFDSVIAR- 472
             |  |  :||:|   :|| |: : :| | ||  ::       :|:: :|||||
Db    604 NKYKSDFMNYQL------HAQMEMLTRKVV---QYMNKYPDNAEIKKI---FESDMKRT 650

Qy    473 -----GHTEPD--------PTQDTELELDGQQVVVPQGQPVPCPEFS 506
                  ||:||         | || | ||  || || ||| ||| ||
Db    651 KEDNYGSLENDALKGYFEKYFLTPFNKIKQIVDDLDKKVEQDQPAPIPENS 701
```

RESULT 8
US-08-910-925-3
; Sequence 3, Application US/08910925
; Patent No. 6162601
GENERAL INFORMATION:
;   APPLICANT: Bandman, Olga
;   APPLICANT: Lal, Preeti
;   APPLICANT: Shah, Purvi
;   TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Incyte Pharmaceuticals, Inc.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/910,925
;   FILING DATE: Herewith
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:

```
Qy    133 --------TKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPV---RTVT 177
                      :|| :| || :| ||:|||| ||:| |  |:|||    ||:|
Db    380 TPKKRLKRDLAANENNQQKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDFSDLL 436

Qy    178 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 237
             ||     |   |   ||   :|     ||:| ::| :: ||:| :
Db    437 TKYNPSVSDRISTNYKTN------TDNHKIAEITIKNL---KLNESQTV---------- 476

Qy    238 EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQA 297
             |:|   |   ||  :|  || |   ||  || |:|  |||   | |   :|:
Db    477 -TLKAKDDSGNVVEKT----FTITVQKKEEKQVP--KTPE----QKD----------- 512

Qy    298 LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQ-LLDSGAPEYKVIQTYLEQTGSNHRCPTL 355
             | | ||:  |  :|  ||| || || |:|   |  :|| :|| |::|:|: |  |::|
Db    513 ------SKTEEKVPQEPKSND----KNQLQELIKSAQQELEKLEKAIKEL-------- 552

Qy    356 QHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYF 415
             ::  :|  :|: :: :|||  |: | :  ||   || : | ::
Db    553 -------MEQPEIPSNPEYGIQKSIWESQKEPIQEAI-TSFKKIIGDSSSKYYTEHYF 602

Qy    416 ASENSKSAGYVIGMKCGAHHVGYMFLGEV-ALGRE--HHIN-TDNPSLKSPPPGFDSVI 470
             |  |  :||:|   :|| |: : :| |   ||||       :|:: :|||||
Db    603 AR------GHTEPD---------NKYKSHFMNYQLHAQMEMLTRKVVQYMNKYPDNAEIKKI---FESDM 646

Qy    471 AR------GHTEPD---------PTQDTELELDGQQVVVPQGQPVPCPEFS 506
                                   | || | ||  || || ||| ||| ||
Db    647 KRTKEDNYGSLENDALKGYFEKYFLTPFNKIKQIVDDFDKKVEQDQPAPIPENS 700
```

RESULT 7
US-08-923-992A-6
; Sequence 6, Application US/08923992A
; Patent No. 6280738
GENERAL INFORMATION:
;   APPLICANT: Tai, Joseph Y.
;   APPLICANT: Blake, Milan S.
;   TITLE OF INVENTION: NO. 6280738-IgA Fc Binding Forms of the Group B
;   TITLE OF INVENTION: Streptococcal Beta Antigens
;   NUMBER OF SEQUENCES: 34
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;   STREET: 1100 New York Avenue, N.W., Suite 600
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20005
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/923,992A
;   FILING DATE: 05-SEP-1997
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/024,707
;   FILING DATE: 06-SEP-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Esmond, Robert W.
;   REGISTRATION NUMBER: 32,893
;   REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 371-2600
;   TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1128 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-923-992A-6

```
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0365 US
; TELECOMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 743 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1684847
US-08-910-925-3
```

```
Query Match        3.9%;    Score 109;  DB 4;  Length 743;
Best Local Similarity  20.0%;  Pred. No. 0.046;
Matches  115;  Conservative   79;  Mismatches  190;  Indels  190;  Gaps   29;

Qy    13 GPEKKKGRQAGREEDP---FRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNC   68
            |  |  | |    :| | ||     |  | |  |   : :| |
Db    86 GGERRTRRESRQESDPEDDDVKKPALQSSVVATSKERTRD------------LIQDQN-  132

Qy    69 TLNQTNIENNNKFYII----QLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKK  123
            ::|  ::|  |   | |:|   | | | :| : |:| :|  |:|
Db   133 -MDEKGKQRNRRIFGLLMGTLQKFKQEST---VATERQNRRQEIEQ-KLE-VQAEEERK  185

Qy   124 DFEKFKFREKTKNNWAERDHFVSHPGKYTLIEVQAE-DEAQEAVVKVDRGPVRTVTKR---  179
            | |: |  |  |   | : |||    | :  ||  | ::|
Db   186 QVENERRELFEERRAKQTELRLLEQKVELAQLQEEWNEHNAKIIKY----IRTKTKPHLF  241

Qy   180 VQPCSLDPATQKLITNIFSKEMFKNTMALMD------LDVKKMPLG------KLSKQQI  226
            : ||||||        |     |        |
Db   242 YIPGRMCPATQKLI----EESQRKMNALFDGRRIEFAEQINKMEARPRRQSMKEKEHQV  296

Qy   227 ARGFEALEALEEALKGPTDGGQSLEEL----SSHFYTVIPHNFGHSQPPPINSPELLQA-  281
            | | |  |  |  ||   |||    |  |  :  ||:|:|
Db   297 VRNEEHKAEQEEG-----KVAQREEELVETGNQH------------------NDVEIEEAG  334

Qy   282 ---KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKV  338
            || :|: |::|  ||        : |  | ||  | |  |:|  |  | |:|
Db   335 EEEEKEIGIVHSDAEKEQ------EEEEQKQEMEVKMEEETEVRESEKQ-QDSQPEEVMD  387

Qy   339 IQTYLEQTGSNHRCPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVRAILTSG  398
            ::|:|        |   :|
Db   388 VLEMVE---------------NVKHVIADQEVMETNRVES------------------  412

Qy   399 LRIMPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPS  458
                           ||   |  |
Db   413 --VEP----------SENEASK------------ELEPEMEFEIEPDK-E  437

Qy   459 LKSPPPGFDSVIA----------RGHTEPDP-----TQDTELELDGQQVVVPQGQP  499
           || | |  |             ||||     ||         :| | |: |
Db   438 CKSLSPGKENVSALDMEKESDEKEKESEPQPEPVAQPAQSQPQLQLQSQSEPQPQLQP  497

Qy   500 VPC-PEFSSSTFSQSEYLIYQESQCRLRYLLEVH  532
            : |    |     ::| | |||:   : :||
Db   498 EPAQPQL------QSPQLQLQSQCHA--VLQSH  523
```

```
RESULT        9
US-08-694-869-1
; Sequence 1, Application US/08694869
; Patent No. 5994123
; GENERAL INFORMATION:
; APPLICANT: Olszewski, N.
; APPLICANT: Tzafrir, I.
; APPLICANT: Somers, D. A.
; APPLICANT: Lockhart, B.
```

```
; APPLICANT: Torbert, K.
; TITLE OF INVENTION: SUGARCANE BACILLIFORM VIRUS
; TITLE OF INVENTION: PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,869
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.369US1
; TELECOMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1871 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-869-1
```

```
Query Match        3.8%;    Score 108.5;  DB 2;  Length 1871;
Best Local Similarity  21.0%;  Pred. No. 0.25;
Matches  100;  Conservative   64;  Mismatches  135;  Indels  177;  Gaps   25;

Qy    33 AEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNNKFYIIQLLQDSN   92
            |||:| |||   |   :| |
Db   919 AEAVK--PPEKK------------SNYELLAKQLLIENSKLKMEKEILIEELN  957

Qy    93 RFFTCWNRWGR-------VGEVGQSKINHFTRLEDAKKDFEKKFREKTKNNWA--------  138
            : | || | :  |    || |   | |||
Db   958 KEIKAHQETKKGKELYIEEASTEVENEIETWKSRAELFEALYNEEVKKNKASTSSVTEGM 1017

Qy   139 ---ERDHFVSHPGKYTLIEVQA------------KVDRGPVRTVTKR---  179
            | |    | | |:          ::||
Db  1018 YQVQIDHL-----RKELREVEATLEVNKVEESEEEAEEVMMASAVKDEMYRFPVIIEVPE 1072

Qy   180 VQPCSLDP-----ATQKLITNIFSKEMF----------KNTMALMDLDVK-  214
            :| | |        |:||| : |         | | |: |
Db  1073 VGKVQLTALLDTGATRSCINQVFIEEKFLQPTKFKVKIHGVNSVTKLDRQVKDGAKLWAG 1132

Qy   215 ---KMPL---GKL---SKQQIARGFEALEALEALEEALKGPTDGGQSLEELSSHFYTVIPHN  264
            | |  | |   ::
Db  1133 ENWFRLPITYVGPMYMGEKTQMLIGCNFMQSLA---------GGVRLEGRTVTFYKYI---  1181

Qy   265 FGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQE--------KTVEEVPH  313
            | |       : |       |:|           :||||
Db  1182 ------ASIKANEYLQAEAEEILV------------ATSEQEFINRSFMSKNKRLLEEMK- 1223

Qy   314 PLDRDY----------QLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVN  362
            || |:        ||::|| |    ||:| :  : |
Db  1224 -EQGYMGEDFLAHWNKNQIKCKIELRN---PDLII----KDKPQTLLNIQK--  1266
```

GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chiu, Choi Ying
; APPLICANT: Duhl, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
; FILE REFERENCE: 200130.442
; CURRENT APPLICATION NUMBER: US/09/172,422A
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-172-422-1

Query Match   3.8%;   Score 108.5;   DB 4;   Length 2548;
Best Local Similarity 21.0%;   Pred. No. 0.42;
Matches   88;   Conservative   63;   Mismatches 126;   Indels 143;   Gaps   25;

Qy   34 EALKAIPAEK-RIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQDSN 92
           ||: ||| :|:|:|| | ||| :::| || :::: |||:||||||||||
Db   757 EILQRCKEEKYSITRKNPRTPLSDLQG----------------MNALNEKNQHDTFDI- 798

Qy   93 RFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDH-------FV 144
           |: | || :|: |||:: :::: || |:: | :|||:| ||
Db   799 ----AWN--GRTG-IRQSRLSSGTSLLDK---DGIFANSTSSKLLERAHGILTRNKNFK 847

Qy   145 SHPG--KYTLIEVQA-----------------EDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQ- 190
           :| |||| || | |||||||:||| | || |
Db   848 SKPALPKH-LLEVNSLKHLTRLTLQDRITKSLLHLH------KKKKPPSISAQFQASL 898

Qy   191 -KLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE--EALKGPTDGG 247
           || :| |::| :|| ::| || |:| | | | :| ||||:||| ||
Db   899 SKLMETLGQAEPY--FVKCIRSNAEKLPL-RFSDVLVLRQLRYTGMLETVYQIRQSGYSSK 955

Qy   248 QSLEELSSHFYTVIPHNFGHSQPPPIN------------SPELLQAKKDMLLVLADIELAQA 297
            ::| ||||||||||||:|| || | || ||||:|||:|||| |||
Db   956 YSFQDFVSHFHVLLPRNI--IPSKFNIQDFFRKINLNPDNYQVGKTMVF------------1002

Qy   298 LQAVSEQEKT-VEEVPHP--------------------LDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS 348
           || ||||| ||||||| || :| ||:|| ||||||||||||||:|||||||
Db 1003 --LKEQERQHLQDLLHQEVLRRIILQRWFRVLLCRQHFL-----------HLRQAS- 1046

Qy   349 NHRCPTLQHIWK--VNQEG------EEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLR 400
            |: ||| || ||::| |||||||:|||||||||||||:|||||||||||||
Db 1047 -----VIIQRFWRNYLNQKQVRDAAVQKDAFV-------------MASAAALLQASWR 1086

RESULT 12
US-08-923-992A-10
; Sequence 10, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005

Qy   363 QEGEEDRFQAH-SKLGNRKLLW------HGTNMAVVAAILTSGLRIMPHSGGRV-GK 411
           :| | | :||||||||||| ||||||||||||||||:||||::|:| ||
Db 1267 ---KEAMRKHIDALLERKVIRPSKSPHRTN---AFIVESGTSIDPKTGKEIRGK 1314

RESULT 10
US-09-349-546-1
; Sequence 1, Application US/09349546
; Patent No. 6093569
; GENERAL INFORMATION:
; APPLICANT: Olszewski, N.
; APPLICANT: Tzafrir, I.
; APPLICANT: Somers, D.A.
; APPLICANT: Lockhart, B.
; APPLICANT: Torbert, K.
; TITLE OF INVENTION: Sugarcane bacilliform virus promoter
; FILE REFERENCE: 600.369US2
; CURRENT APPLICATION NUMBER: US/09/349,546
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: US 08/694,869
; EARLIER FILING DATE: 1996-08-09
; EARLIER APPLICATION NUMBER: PCT/IB97/01338
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1871
; TYPE: PRT
; ORGANISM: sugarcane bacilliform virus
US-09-349-546-1

Query Match   3.8%;   Score 108.5;   DB 3;   Length 1871;
Best Local Similarity 21.0%;   Pred. No. 0.25;
Matches 100;   Conservative   64;   Mismatches 135;   Indels 177;   Gaps   25;

Qy   33 AEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQDSN 92
           |||| | |:| :::::| | :::| |::
Db   919 AEAVK--PPEKK----------------SNYELLAKQLLIENSKLKMEKEILIEELN 957

Qy   93 RFFTCWNRWGR-----VGEVGQSKINHFTRLEDAKKDFEKKFREKTKNNWA---------- 138
           | | | | || ||| |||:|| ||||||:||| | :|:|||
Db   958 KEIKAHQETKKGKELYIEEASTEVENEIETWKSRAELFEALYNEEVKKNKASTSSVTEGM 1017

Qy   139 ---ERDHFVSHPGKYTLIEVQA------EDEAQEAVV-----KVDRGPVRTVTKR 179
             :| | :| || | |: :| |||| || ||||||||||:|||
Db 1018 YQVQIDHL------RKELREVEATLEVNKVEESEEEAEEVMASAVKDEMYRFPVIIEVPE 1072

Qy   180 VQPCSLDP----ATQKLITNIFSKEMF-----------KNTMALMDLDVK------ 214
           : :: :: ||::||||||||:|| |:|:|||||||||
Db 1073 VGKVQLTALLDTGATRSCINQVFIEEKFLQPTKFKVIHGVNSVTKLDRQVKDGAKLWAG 1132

Qy   215 ----KMPL---GKL----SKQQIARGFEALEALEEEALKGPTDGGQSLEELSSHFYTVIPHN 264
           :||: ||: ||| :||||||:|||| ||| :| :|||:|||||||||||:
Db 1133 ENWFRLPTYVGPMYMGEKTQMLIGCNFMQSLA---------GVRLEGRTVTFYKYI- 1181

Qy   265 FGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQE------------KTVEEVPH 313
            |||| || || || | |:||:||:|||| | |:|| ||||||||
Db 1182 ------ASIKANEYLQAEAEEILV--------ATSEQEFINRSFMSKNKRLLEEMK 1223

Qy   314 PLDRDY--------QLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVN 362
             : |: ||||||||||||||||||||:||||||||||:|||||||:||
Db 1224 --EQGYMGEDTLAHWKNQIKCKIELRN---KDKPQTLLNIQK-- 1266

Qy   363 QEGEEDRFQAH-SKLGNRKLLW------HGTNMAVVAAILTSGLRIMPHSGGRV-GK 411
           :| | | :||||||||||| ||||||||||||||||:||||::|:| ||
Db 1267 ---KEAMRKHIDALLERKVIRPSKSPHRTN---AFIVESGTSIDPKTGKEIRGK 1314

RESULT 11
US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485

```
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/923,992A
;     FILING DATE: 05-SEP-1997
;     CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 60/024,707
;     FILING DATE: 06-SEP-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Esmond, Robert W.
;     REGISTRATION NUMBER: 32,893
;     REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (202) 371-2600
;     TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 10:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1164 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;US-08-923-992A-10


Query Match             3.8%;  Score 107.5;  DB 4;  Length 1164;
Best Local Similarity  20.4%;  Pred. No. 0.14;
Matches  108;  Conservative  68;  Mismatches 210;  Indels 143;  Gaps  24;

Qy   21 QAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN 79
         ||:|||::|||:|:|||
Db  309 QATQVKNQFLENAQKLKEIQ----------PLIKETNVKLYKAMSESLEQVEKQLKHN 356

Qy   80 NKFYIIQLLQDSNRFFTCW----NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFR--EK 132
         :::|||:| :| :|:| |: |: |||| |
Db  357 SQANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFKTSEQV 416

Qy  133 TKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKL 192
         :|||  | ||  :||  ::|
Db  417 TPKKRVKRD-LAANENNQQKIELTVSPEN----ITVYEG--------------EDL 453

Qy  193 ITNIFSKEMFKNTMALMDLDVKKMP---------LGKLSKQQIA----RGFEALEALEE 238
         :||  :|| ||
Db  454 KFTLTAKSDSKTTLDFSDLLTKYNPSVSDRISTNYKTNTDNHKIAEITIKNLKLNESQTV 513

Qy  239 ALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQAL 298
         ||  |  |:: ||| |::: || |:| ||
Db  514 TLKAKDDSGNVVQKT---FTITVQKEEKQVP-KTPE----QKD---------------- 549

Qy  299 QAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQH 357
         | ||::|::|| || | ||:|| ||| : |
Db  550 ------SKTEEKVPQEPKSND----KNQLQEL------IKSAQQQLEKLEK-------- 584

Qy  358 IWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFAS 417
         || ||:  : | ||  :|  |:|  ||
Db  585 --AIKELMEQEPIPSNPEYGIQKSTWESQKEPIQEAI-TSFKKIIGDSSSKYYTEHYFNK 641

Qy  418 ENSKSSAGYVIGMKCGAHHVGYMFLGEVALGREHHIN--TDNPSLKSPPPGFDSVIAR-- 472
         || |:  || |: ||  |:|| ::
Db  642 YKSDEFMNYQL-----HAQMEMLTRKVV----QYINKYPDNAEIKKI--FESDMKRTKE 688

Qy  473 --GHTEPD---------PTQDTELELDGQQVVVPGQPVPCPEFS 506
         | || |  |: | |
Db  689 DNYGSLENDALKGYFEKYFLTPFNKIKQIVDDLDKKVEQDQPAPIPENS 737

RESULT 13
US-09-045-360-2
; Sequence 2, Application US/09045360
; Patent No. 6207880
```

```
;   GENERAL INFORMATION:
;     APPLICANT: Jens Kossmann
;     APPLICANT: Ruth Lorberth
;     TITLE OF INVENTION: PLANTS WHICH SYNTHESIZE A MODIFIED STARCH,
;     TITLE OF INVENTION: PROCESS FOR THE PRODUCTION THEREOF AND MODIFIED STARCH
;     NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: FISH & NEAVE
;     STREET: 1251 Avenue of the Americas
;     CITY: New York
;     STATE: New York
;     COUNTRY: USA
;     ZIP: 10020
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/045,360
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PCT/EP96/04109
;     FILING DATE: 19-SEP-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: DE 19547733.2
;     FILING DATE: 20-DEC-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: DE 19534759.5
;     FILING DATE: 19-SEP-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Haley Jr., James F.
;     REGISTRATION NUMBER: 27,794
;     REFERENCE/DOCKET NUMBER: GFB-5
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 212-596-9000
;     TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1464 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;US-09-045-360-2


Query Match             3.7%;   Score 105.5;  DB 4;  Length 1464;
Best Local Similarity  20.8%;   Pred. NO. 0.33;
Matches   91;  Conservative  73;  Mismatches 193;  Indels  81;  Gaps  18;

Qy   37 KAIPAE-KRIIRVDPTCPLSSNPGTQVYEDYNCTLN---QTNIE---NNNNKFYIIQLLQ 89
         | ||| ||
Db   61 KKIPMEKKRAFSSSPHAVLTTDTSSELAEKFSLGGNIELQVDVRPPTSGDVSFVDFQVTN 120

Qy   90 DSNRFFTCWNRWGRVGEVGQSKINHFT-RLEDAKKDFEKKFREKTKNNWAERDHFVSHPG 148
         | : || |: ::
Db  121 GSDKLFLHW---------GAVKFGKETWSLPNDRPDGTKVYKNK----ALRTPFVK-SG 165

Qy  149 KYTLIEVQAEDEAQEAV---------VKVDRGPVRTVTKRVQPCSLDPATQKLLITNI 196
         :|
Db  166 SNSILRLEIRDTAIEAIEFLIYDEAHDKWIKNGGNFRVKLSRKEIRGPDVSVPEELVQI 225

Qy  197 FS-------KEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAALKGPTDGGQ 248
         :|||| :
Db  226 QSYLRWERKGKQNYPPEKEKEYEAARTVL---QEEIARG-ASIQDIRARLTKTNDKSQ 280

Qy  249 SLEELSSHFYTVIPHNFGHSQ---------PPPINSPELLQAKKDMLLVL-ADIEL 294
         : :|: :| || |
Db  281 SKEEPLHVTKSDIPDDLAQAQAYIRWEKAGKPNYPPEKQIEELEEARRELQLELEKGITL 340

Qy  295 AQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSG--------APEYKVIQTYLEQT 346
         :| | | || :| | :| | |
```

```
Db  168 TLKV---------ALAKKEVEAKELEIEKLQYEISTLEQEVATAQHQVDNLKKLLAG 215
                     : :: :::  :: :     :  : :::  :
Qy  243 --PTDGGSLEELSSHFVTVIPINFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA 300
                :  :  :: ::: :::         ::  :::
Db  216 ADPDDGTEVIEAKLKK------GEAE--LNAKQAELAKK----QTELEKLLDS 256
                    ::  :    :  : ::  :::
Qy  301 VSEQEKTVEEVPHPLDRDYQLLKCQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 360
                  ::  ::  ::::   ::  ::    ::
Db  257 LDPEGKTQDE----LDKEAEEAE-----LDKKADELQNKVADLEKEISN-----LEILL 301
                 ::   :  :: :: :::
Qy  361 VNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420
                                    :
Db  302 GGADXEDDTAALQNKLATKKAELEKTQKELDDAALNELG----PD----------- 341
Qy  421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480
                      : :  :  : : :
Db  342 ---------------GDEEETPAPAPQPEQPAPA----------------PKPE 364
Qy  481 QDTELELDGQQVVPQ-GQPVPCPE 504
                 :: :  :: :::: :
Db  365 QPAPAPKPEQPAPAPKPEQPAPAPK 389
```

```
RESULT 14
US-08-961-083-160
; Sequence 160, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
;  APPLICANT: Choi et. al.
;  TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;  NUMBER OF SEQUENCES: 452
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,083
;   FILING DATE:
;  PRIOR APPLICATION DATA:
;   CLASSIFICATION: 435
;   LENGTH: 641 amino acids
;   APPLICATION NUMBER:
;   FILING DATE:
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P2
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
;  MOLECULE TYPE: protein
US-08-961-083-160
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:

Query Match             3.7%;  Score 105; DB 4; Length 641;
Best Local Similarity  19.0%;  Pred. No. 0.092;
Matches  96;  Conservative  67;  Mismatches 168;  Indels 174;  Gaps  21;

Qy   16 KKKGRQAGR--EEDFPRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQT 73
            ||| |   :     |:|    |::::            ||  :
Db   43 KKKAEDAQKKYEDDQKRTEEKARKEAEASQKLNDVALVV------QNAYKEYREVQNQR 95
            |                    : :::: :     |
Qy   74 NIENNNNKFYIQLLQDSNRFFTCWNRWGRVGEVGQSKINHF-TRLEDAKK--DFEKKF 129
               : :::                       :|:| | :|: | : | ||
Db   96 SKYKSDAEV--------------------------QKKLTEVDSKIEKARKEQQDLQNKF 129
             : ::: :
Qy  130 REKTKNNWAERDHFVSHPGKYTLIEVQAED-EAQEAVVKVDRGPVRTVTKRVQPCSLDPA 188
         |    :    ::|| | ||   :           ||
Db  130 NE------VRAVVVPEPNALAETKKKAEEAKAEEKVAK----------RKYDYA 167
            |:           | : ::
Qy  189 TQKLITNIFSKEMFKNTWALMDLVKKMPLGKLS-----KQQIARGFEALEALEEALKG 242
           |:              || :  ::  :| :  || ||  ::    |
```

```
RESULT 15
US-08-326-117B-2
; Sequence 2, Application US/08326117B
; Patent No. 5693491
; GENERAL INFORMATION:
;  APPLICANT: BULLA, LEE A.
;  APPLICANT: JI, TAE
;  TITLE OF INVENTION:  RECEPTOR FOR A BACILLUS THURINGIENSIS
;  TITLE OF INVENTION:  TOXIN
;  NUMBER OF SEQUENCES: 26
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORRISON & FOERSTER
;   STREET: 2000 Pennsylvania Ave. N.W.
;   CITY: Washington
;   STATE: DC
;   COUNTRY: USA
;   ZIP: 20006-1812
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/326,117B
;   FILING DATE: 19-OCT-1994
;   CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;   NAME: MILLMAN, ROBERT A
;   REGISTRATION NUMBER: 36,217
;   REFERENCE/DOCKET NUMBER: 7112-0037.00
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 887-1500
;   TELEFAX: (202) 887-0763
;   TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1528 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;  MOLECULE TYPE: protein
US-08-326-117B-2

Query Match             3.7%;  Score 103.5; DB 1; Length 1528;
Best Local Similarity  20.3%;  Pred. No. 0.57;
Matches 120;  Conservative  67;  Mismatches 207;  Indels 197;  Gaps  31;

Qy   58 PGTQVYEDYN----------CTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGE 106
            : : |:|:           :::: ||:  | |||||
Db  142 PGMQQYM-FNVRVDGQSLVAGVSLAIVNIDDNAP---IIQ------NFEPC------RVPE 186
           || ||          :: || :|
```

```
Qy   107  VGQSKINHFT----------------RLEDAKKDFEKKFREKTK--NNWAERD-------  141
          :|:.|.          ||:.:.|.::|.| ::|:         :
Db   187  LGEPGLTECTYQVSDADGRISTEMTFRIDSVRGDEETFYIERTNIPNQWMWLNMTIGVN  246

Qy   142  ---HFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFS  198
          :||:.|           |.|         :.:.||:
Db   247  TSLNFVTSP--LHIFSVTALDSLPN-------THTVTMVQVANVNSRPPRWL-EIFA  294

Qy   199  KEMFKN------TMALMDLDVK-KMPLGKLSKQQIARGFEALEEALKG------PTD  245
          :.:  |:.:|.|:  ||:       |.:.:|||
Db   295  VQQFEEKSYQNFTVRAIDGDTEINMPINYRLITNEEDTFFSIEALPGGKSGAVFLVSPID  354

Qy   246  GGQSLEEL----------------SSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVL  289
          .:                |.|.:  :.:||.|     ::.|
Db   355  RDTLQREVFPLTIVAYKYDEEAFSTSTNVVIIVTDINDQRPEPIHKEYRLAIMEETPLTL  414

Qy   290  ------ADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQ-----------LL  329
          .:|             |.        |        .:|
Db   415  NFDKEFGFHDKDLGQNAQYTVRLESV--DPPGAAEAFYIAPEVGYQRQTFIMGTLNHSML  472

Qy   330  DSGAPEYKVIQTYLEQTGSN--HRCPTLQHIWKVNQEGEEDRRQ--------------  371
          |.|||:.|:.:.:|:|     ||::  |:.||.:|
Db   473  DYEVPEFQSITIRVVATDNNDTRHVGVALVHIDLINWNDEQPIFEHAVQTVTFDETEGEG  532

Qy   372  -------AHSK-LGN--RKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS  420
          .|         |.      |.        |:  .:|:
Db   533  FFVAKAVAHDRDIGDVVEHTLLGNAVNFLTIDK-LTGDIRV-------------SAND  576

Qy   421  KSAGYVIGMKCGAHHVGYMFLGEVA---LGREHH----------INTDNPSLKSP--P  463
          .:    |     |:  :.|.|        |:|::.|.|
Db   577  -SFNY-------HRESELFVQVRATDTLGEPFHTATSQLVIRLNDINNTPPTLRLPRGS  627

Qy   464  PGFDSVIARGH-------TEPDPTQDTELELD-GQQVVVPQGQPVPCPEF  505
          |.|    :||   :.|||:.      ||
Db   628  PQVEENVPDGHVITQELRATDPDTTADLRFEINWDTSFATKQGRQANPDEF  678
```

Search completed: August 29, 2002, 07:58:13
Job time: 290 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:02:11 ; Search time 36.19 Seconds
(without alignments)
570.255 Million cell updates/sec

Title: US-09-701-586B-4
Perfect score: 2823
Sequence: 1 MAPKPKPWVQTEGPEKKKGR..........EYLIYQESQCRLRYLLEVHL 533

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SUMMARIES

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2811	99.6	533	1	PPO3_HUMAN	Q9y6f1 homo sapien
2	738	26.1	559	1	PPO2_MOUSE	O88554 mus musculu
3	729.5	25.8	583	1	PPO2_HUMAN	Q9ugn5 homo sapien
4	688.5	24.4	637	1	PPOL_ARATH	Q11207 arabidopsis
5	677	24.0	1015	1	PPOL_BOVIN	P18493 bos taurus
6	674.5	23.9	996	1	PPOL_SARPE	P21208 sarcophaga
7	674.5	23.9	1011	1	PPOL_CHICK	P26446 gallus gall
8	673.5	23.9	1012	1	PPOL_CRIGR	Q9r152 cricetulus
9	666.5	23.6	1013	1	PPOL_HUMAN	P09874 homo sapien
10	665.5	23.6	1013	1	PPOL_RAT	P27008 rattus norv
11	653	23.1	998	1	PPOL_XENLA	P31669 xenopus lae
12	649.5	23.0	1012	1	PPOL_MOUSE	P11103 mus musculu
13	638	22.6	994	1	PPOL_DROME	P35875 drosophila
14	393.5	13.9	538	1	YQN4_CAEEL	Q09525 caenorhabdi
15	258	9.1	1724	1	PPOV_HUMAN	Q9ukk3 homo sapien
16	190.5	6.7	135	1	PPOL_ONCMA	Q08824 oncorhynchu
17	126	4.5	1222	1	YMP3_CAEEL	Q10947 caenorhabdi
18	120	4.3	1938	1	MYS_AEQIR	P24733 aequipecten
19	115.5	4.1	1395	1	SP41_YEAST	P38904 saccharomyc
20	114	4.0	918	1	HXK1_BOVIN	P27595 bos taurus
21	113.5	4.0	1164	1	BAG_STRAG	P27951 streptococc
22	113.5	4.0	4540	1	DYHC_PARTE	Q27171 paramecium
23	109.5	3.9	446	1	GAG_OMVVS	P16900 ovine lenti
24	108.5	3.8	382	1	MYB_AVIMB	P01104 avian myelo
25	107	3.8	1186	1	SMC_BACSU	P51834 bacillus su
26	106	3.8	724	1	HMMR_HUMAN	O75330 homo sapien
27	106	3.8	932	1	YAIA_SCHPO	Q09897 schizosacch
28	105.5	3.7	2025	1	TTC3_HUMAN	P53804 homo sapien
29	105	3.7	882	1	RA50_PYRFU	P58301 pyrococcus
30	105	3.7	1057	1	POL_SIVAI	Q02836 simian immu
31	104.5	3.7	641	1	MYB_CHICK	P01103 gallus gall
32	104.5	3.7	1403	1	VG22_HSVI1	Q00105 ictalurid h
33	104	3.7	1324	1	CUT3_SCHPO	P41004 schizosacch
34	104	3.7	1433	1	REST_CHICK	O42184 gallus gall
35	103	3.6	716	1	RRP2_IALE2	P26123 influenza a
36	103	3.6	1940	1	MYH3_RAT	P12847 rattus norv
37	102.5	3.6	640	1	MYB_HUMAN	P10242 homo sapien
38	102.5	3.6	704	1	VPS1_YEAST	P21576 saccharomyc
39	102.5	3.6	3210	1	CENF_HUMAN	P49454 homo sapien
40	102	3.6	716	1	RRP2_IALE1	P26122 influenza a
41	101.5	3.6	584	1	LIGA_HUMAN	P41214 homo sapien
42	101.5	3.6	612	1	EXO2_BPT5	P11109 bacterlopha
43	101.5	3.6	874	1	SYA_PASMU	P57933 pasteurella
44	101.5	3.6	3075	1	LMA1_HUMAN	P25391 homo sapien
45	101	3.6	782	1	L100_ADE12	P36714 human adeno

ALIGNMENTS

RESULT 1
PPO3_HUMAN
ID   PPO3_HUMAN   STANDARD;   PRT;   533 AA.
AC   Q9Y6F1; Q9UG81;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+) ADP-
DE   ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (pADPRT-3)
DE   (hPARP-3).
GN   ADPRTL3 OR PARP3 OR ADPRT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=9263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [2]
RP   SEQUENCE OF 75-533 FROM N.A.
RC   TISSUE=Kidney;
RA   Ansorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels are in
CC   the kidney, skeletal muscle, liver, heart and spleen; also
CC   detected in pancreas, lung, placenta, brain, leukocytes, colon,
CC   small intestine, ovary, testis, prostate and thymus.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF083068; AAD29855.1; -.
DR   EMBL; AL050034; CAB43246.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation.
FT   DOMAIN   14   20   NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   CONFLICT   80   80   K -> N (IN REF. 2).

```
FT   CONFLICT      171     171          A -> G (IN REF. 2).
FT   CONFLICT      411     411          K -> E (IN REF. 2).
SQ   SEQUENCE      533 AA;  60117 MW;  7C0AB89E64D1B9FD CRC64;

Query Match             99.6%;  Score 2811;  DB 1;  Length 533;
Best Local Similarity   99.6%;  Pred. No. 5.1e-189;
Matches  531;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy      1 MAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60

Qy     61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120

Qy    121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 180
            ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRAPVRTVTKRV 180

Qy    181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240

Qy    241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 300

Qy    301 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 360

Qy    361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420

Qy    421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480

Qy    481 QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 533
```

```
RESULT   2
ID   PPO2_MOUSE   STANDARD;   PRT;   559 AA.
AC   O88554; Q99N29;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE   (mPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
```
```
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Embryo;
RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129Sv;
RX   MEDLINE=21179160; PubMed=11133988;
RA   Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,
RA   Niedergang C.P.;
RT   "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2
RT   (PARP-2) gene to the gene for RNase P RNA.";
RL   J. Biol. Chem. 276:11092-11099(2001).
RN   [3]
RP   SEQUENCE OF 9-559 FROM N.A.
RC   STRAIN=129/Sv X C57BL/6;
RX   MEDLINE=99268466; PubMed=10038144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
```
```
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC       SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in
CC       testis followed by ovary.
CC   -!- INDUCTION: By high levels of DNA-damaging agents.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ007780; CAA07679.1; -.
DR   EMBL; AF191547; AAK13253.1; -.
DR   EMBL; AF072521; AAC25415.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1341112; Adprt2.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND      1      65               POTENTIAL.
FT   DOMAIN       66     559               NAD-BINDING (BY SIMILARITY).
FT   DOMAIN        3       9               NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                         (POTENTIAL).
FT   DOMAIN       33      39               NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                         (POTENTIAL).
FT   CONFLICT     82      82               L -> V (IN REF. 2).
FT   CONFLICT    177     177               V -> I (IN REF. 2).
FT   CONFLICT    486     486               R -> Q (IN REF. 2).
SQ   SEQUENCE    559 AA;  63396 MW;  E0AEDAEE412C1445 CRC64;
```

```
Query Match             26.1%;  Score 738;  DB 1;  Length 559;
Best Local Similarity   34.3%;  Pred. No. 3.5e-44;
Matches  187;  Conservative  103;  Mismatches  191;  Indels  64;  Gaps  19;

Qy     16 KKKGRQA-GREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVY---EDYNCTL 70
            ::||  | ||  |:| :::  :::|:| : ::  |:   | |    |  ||  : :  |
Db     42 QRKGPMAGGKDADRTKDNRDSVKTLLLKGK-APVDPECAAKLGK-AHYCEGDDVYDVML 99

Qy     71 NQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDAKKDFEKK 128
            ||||||||||||||||| ||||||||||||||||||||| |||||||||||||||||||
Db    100 NQTNLQFNNNKYTLIQLLEDDAQRNFSVMWRWGRVGKTGQHSLVTCSGDLNKAKEIFQKK 159

Qy    129 FREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEA---QEAVVKVDRGPVRTVTKRVQ 181
            :  |||||  :| |||:  |  |:||    :|   :    :| | |:| : | :  :|:
Db    160 FLDKTKNNWEDRENFEKVPGKYDMLQMDYAASTQDESKTKEETLKPE------------ 207

Qy    182 PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 241
```

```
Db  208  -SQLDLRVQELLKLICNVQTMEEMIEMKYDTKRAPLGKLTVAQIKAGYQSLKKIEDCIR  266
Qy  242  GPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVIADIELAQAIQAV  301
Db  267  A-GQHGRALVEACNEFYTRIPHDFGLSIPPVIRTEKELSDKVKLLEALGDIEI--ALKLV  323
Qy  302  SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHR--CPTLQHI  358
Db  324  KSERQGLE---HPLDQHYRNLHCALRPLDHESNEFKVISQYLQSTHAPTHKDYTMTLLDV  380
Qy  359  WKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVCKGIY  414
Db  381  FEVEKEGEKEAFR--EDLPNRMLLWHGSRLSNWVGIISHGLRVAPPEAPITGYMFCKGIY  438
Qy  415  FASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH  474
Db  439  FADMSSKSANYCFASR--LKNTGLLLSEVALGQCNELLEANPKAQGLLRGKHSTKGMGK  496
Qy  475  TEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSS--------TFSQSEYLIYQESQCRLR  526
Db  497  MAPSPA--------HFITLNGSTVPLGPASDTGILNPEGYTLNYNEFIVTVSPNQVRMR  546
Qy  527  YLLEV  531
Db  547  YLLKI  551
```

```
RESULT 3
PPO2_HUMAN
ID   PPO2_HUMAN     STANDARD;     PRT;    583 AA.
AC   Q9UGN5; Q9Y6C8; Q9NUV2; Q9UMR4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-
DE   2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoeger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE OF 2-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [3]
RP   SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fibroblast;
RX   MEDLINE=99268466; PubMed=10338144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA   Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA   Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA   Masuho Y., Kanehori K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC       SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (BY
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY similarity).
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN
CC       THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO
CC       DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;
CC       LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND
CC       THYMUS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
```

```
DR   EMBL; AJ236912; CAB65088.1; -.
DR   EMBL; AF085734; AAD29857.1; ALT_INIT.
DR   EMBL; AJ236876; CAB41505.2; ALT_INIT.
DR   EMBL; AK001980; BAA92017.1; ALT_TERM.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Alternative splicing.
FT   DNA_BIND      1     88       POTENTIAL.
FT   DOMAIN       86    583       NAD-BINDING (BY SIMILARITY).
FT   DOMAIN        4      7       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                (POTENTIAL).
FT   DOMAIN       35     40       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                (POTENTIAL).
FT   VARSPLIC     68     80       MISSING (IN ISOFORM 2).
FT   CONFLICT    447    447       P -> H (IN REF. 2).
FT   CONFLICT    481    481       N -> H (IN REF. 4).
SQ   SEQUENCE   583 AA;   66205 MW;  5B7AE8AE531836AF CRC64;
```

```
Query Match             25.8%;  Score 729.5;  DB 1;  Length 583;
Best Local Similarity   35.4%;  Pred. No. 1.4e-43;
Matches  201;  Conservative  95;  Mismatches 184;  Indels 87;  Gaps 22;

Qy    15  EKKKGRQAG----------REED------PFRSTA-----EALKAIPAEKRIIRVDPTCPLSSNP   58
Db    46  ESKKMPVAGGKANKDRTEDKQDGMPGRSWASKRVSESVKALLLKGK-APVDPEC--TAKV  102
Qy    59  G-TQVY----EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SK  111
Db   103  GKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSL  162
Qy   112  INHFTRLEDAAKDFEKKFPREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVK  167
Db   163  VACSGNLKAKEIFQKKFLDKTKNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEE  222
Qy   168  VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA  227
Db   223  SLKSPLKPESQ------LDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIK  276
Qy   228  RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLL  287
```

```
DR    SMART; SM00513; SAP; 2.
KW    Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein:
KW    ADP-ribosylation.
FT    DNA_BIND       1    140      POTENTIAL.
FT    DOMAIN         1    140      NAD-BINDING (BY SIMILARITY).
FT    DOMAIN        41     62      NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ    SEQUENCE     637 AA;  72175 MW;  527A8F464605D127 CRC64;

      Query Match             24.4%;  Score 688.5;  DB 1;  Length 637;
      Best Local Similarity   36.2%;  Pred. No. 1.2e-40;
      Matches  174;  Conservative  89;  Mismatches 185;  Indels  33;  Gaps  16;

Qy       66  YNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK 123
                : !!!!:! !!!!:: !!!:::!!:!!: !!!!!: :!:!: !!!!!
Db      169  YDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVYTRWGRVGCVKGQSKLDGPYDSWDRAIE 228

Qy      124  DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVKRVQP- 182
              !:! :!!!! !!!:!:! !!!: !! :!!!!!!  !:::!!!  !:!  : !:!
Db      229  IFTNKFNDKTKNYWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNND---IPSSSSEVKPE 285

Qy      183  -CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 241
                !  : :  :    :!!: :  :! : ! :!!!:! !!: !!! !!::!
Db      286  QSKLLDTRVAKFISLLICNVSMMAQHMEIGYNANKLPLGKISKSTISKGYEVLKRISEVI- 344

Qy      242  GPTDGGQSLEELSSHFYTVIPHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ 299
              : !!!!:!:  :  !!!:!!:   !: !!!!!!!!
Db      345  -DRYDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDPQKLLKQKIEMVEALGEIELATKLL 403

Qy      300  AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ- 356
               !  !!         ! !!! :!! :!
Db      404  SVDPGLQ-----DDPLYYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVEIA 458

Qy      357  HIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKG 412
              !!!!!!!  !!!!!!!!!!!!!!:!!! !!!  !!!!!!!!:::!!
Db      459  QLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSOGLRIAPPEAPVTGYMFGKG 518

Qy      413  IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 472
              :!!! !!!!!!!!! !!!!!!!!!!!!!!!!!! ::!!!!!!!
Db      519  VYFADMFSKSANYCYA-NTGAND-GVLLCEVALGDMNELLYSDYNADNLPPGKLSTKGV 576

Qy      473  GHTEPDTPQDTELELDGQQVVVPQGQPV--PCPEFSSSTFSQSEYLIYQESQCRLRYLLE 530
              ! !!!!!!!!  !!! !!!!!!!  :!!:!!       ::!!:::
Db      577  GKTAPNPSEAQTLE-DG--VVVPLGKPVERSC--SKGMLLYNEYIVYNVEQIKMRYVIQ 630

Qy      531  V 531

Db      631  V 631
```

RESULT 5
```
PPOL_BOVIN
ID    PPOL_BOVIN     STANDARD;     PRT;   1015 AA.
AC    P18493; Q9TS00;
DT    01-NOV-1990 (Rel. 16, Created)
DT    01-NOV-1990 (Rel. 16, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
DE    Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE    ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN    ADPRT.
OS    Bos taurus (Bovine).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC    Bovidae; Bovinae; Bos.
OX    NCBI_TaxID=9913;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=90382673; PubMed=2119324;
RA    Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;
RT    "Cloning of a full-length cDNA encoding bovine thymus
RT    poly(ADP-ribose) synthetase: evolutionarily conserved segments and
RT    their potential functions.";
RL    Gene 90:249-254(1990).
```

Second column (left side):

```
Db      277  AGYQSLKKIEDCIRA-GQHGRALMEACNEFYTRIPHDFGLRTPLIRTQKELSEKIQLLE 335
              :!!::!:  !!: ::         !!! !!!!!!! :!! !
Qy      288  VLADIELAQALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 346
              ! !!!:!!!:!  !         :!!!:! !:!!! !!!!:! !!! !
Db      336  ALGDIEIAIKL------VKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQST 389

Qy      347  GS---NHRCPTLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP 403
                   :!!!!!!!! !!!!!!!!!!!!!!!!!!!!! !::!!!!!!!!!!!!:!!
Db      390  HAPTHSDYTMTLLDLFEVERDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAP 447

Qy      404  H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL 459
              !      !!!!!:! !!!!!!       !:!  !! :!  !!!!!!!!!!
Db      448  PEAPITGYMFGKGIYFADMSSKSANYCFASR--LKNTGLLLLSEVALGQCNELLEANPKA 505

Qy      460  ----------KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVVPQGQPVPCPE 504
                        :!!!!!!!! !!          !! :!
Db      506  EGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPASDT-------GILNPDGY----- 552

Qy      505  FSSSTFSQSEYLIYQESQCRLRYLLEV 531
              !!!!!!:!!!! :!!!!!!!!
Db      553  ---TLNYNEYIVYNPNQVRMRYLLKV 575
```

RESULT 4
```
PPOL_ARATH
ID    PPOL_ARATH     STANDARD;     PRT;   637 AA.
AC    Q11207;
DT    01-OCT-1996 (Rel. 34, Created)
DT    01-OCT-1996 (Rel. 34, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
DE    Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE    ribosyltransferase) (Poly[ADP-ribose] synthetase).
APP.
GN    Arabidopsis thaliana (Mouse-ear cress).
OS    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC    eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX    NCBI_TaxID=3702;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=95269779; PubMed=7750552;
RA    Lepiniec L., Babiychuk E., Kushnir S., van Montagu M., Inze D.;
RT    "Characterization of an Arabidopsis thaliana cDNA homologue to animal
RT    poly(ADP-ribose) polymerase.";
RL    FEBS Lett. 364:103-108(1995).
CC    -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC        PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC        ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC        CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC        TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC        EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC    -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC        nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC    -!- SUBCELLULAR LOCATION: Nuclear.
CC    -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC    --------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute.  There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------------
DR    EMBL; Z48243; CAA88288.1; -.
DR    HSSP; P26446; 1A26.
DR    InterPro; IPR001290; PARP.
DR    InterPro; IPR004102; PARP_reg.
DR    InterPro; IPR003034; SAP.
DR    Pfam; PF00644; PARP; 1.
DR    Pfam; PF02877; PARP_reg; 1.
DR    Pfam; PF02037; SAP; 2.
```

```
RN   [2]
RP   SEQUENCE OF 647-714 AND 838-903 FROM N.A.
RX   MEDLINE=88151954; PubMed=2450019;
RA   Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,
RA   Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;
RT   "Depression in gene expression for poly(ADP-ribose) synthetase during
RT   the interferon-gamma-induced activation process of murine macrophage
RT   tumor cells.";
RL   Eur. J. Biochem. 171:571-575(1988).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   -------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------
DR   EMBL; D90073; BAA14114.1; -.
DR   EMBL; X06986; CAA30046.1; -.
DR   EMBL; X06987; CAA30047.1; -.
DR   PIR; JS0428; JS0428.
DR   PIR; S00328; S00328.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0
FT   DNA_BIND      1    374       BY SIMILARITY.
FT   DOMAIN      375    525       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      386    462       BRCT.
FT   DOMAIN      526   1015       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     127    164       PARP-TYPE.
FT   DOMAIN      209    211       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     408    408       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     414    414       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     446    446       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     449    449       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     457    457       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     472    472       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     489    489       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     492    492       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     514    514       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     515    515       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     521    521       ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE   1015 AA;  113355 MW;  0A5FE9D9F04F5B04 CRC64;

Query Match             24.0%;  Score 677;  DB 1;  Length 1015;
Best Local Similarity   33.8%;  Pred. No. 1.4e-39;
Matches  187;  Conservative   99;  Mismatches  205;  Indels   62;  Gaps  24;

Qy    1 MAPKPK---PWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSS 56
                 :|: ||:  || :.| :|::| |||::|
Db  494 VGPKGKSGAAPSKKSKGPVKEEG--TNKSEKRMKLTLKGGAA-------VDPDSGLEH 542

Qy   57 NPGTQVYED---YNCTLNQTNIENNNNKFYIIQLLQDSNRF-FTCWNRWRGRVGEV-GQS 110
         |--| |   | :| :||:|   | :.|| :-| :
Db  543 N--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVIGSN 600

Qy  111 KINHFTRLEDAKDFEKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVD 169
Db  601 KLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKHPKKFYPLEIDYGQDE--EAVKKLT 657

Qy  170 RGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARG 229
Db  658 VNP-GTKSKLPKP------VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAA 710

Qy  230 FEALEALEAL-KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLV 288
Db  711 YSILSEVQQALSQGSSD--SHILDLSNRFYTLIPHDFGMKKPPLLNNANSVQAKVEMLDN 768

Qy  289 LADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLDSGAPEYKVIQTYLEQT-G 347
Db  769 LLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDKDSEEAEIIRKYVKNTHA 824

Qy  348 SNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH- 404
Db  825 TTHNAYDLEVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTNFAGILSQGLRIAPPE 884

Qy  405 ---SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDNPS 458
Db  885 APVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEAALGMYELKHAR-H 938

Qy  459 LKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIY 518
Db  939 ISKLPKGKHSVKGLGKTTPDPS--ASITVDG--VEVPLGTGI-SSGVNDTCLLYNEYIVY 993

Qy  519 QESQCRLRYLLEV 531
Db  994 DIAQVHLKYLLKL 1006

RESULT 6
PPOL_SARPE
ID   PPOL_SARPE     STANDARD;      PRT;   996 AA.
AC   Q11208;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
OS   Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7386;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=94170813; PubMed=8125121;
RA   Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,
```

```
RA   de Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;
RT   "Cloning and functional expression of poly(ADP-ribose) polymerase
RT   cDNA from Sarcophaga peregrina.";
RL   Eur. J. Biochem. 220:607-614(1994).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D16482; BAA03943.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; FALSE_NEG.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; zinc-finger; Zinc.
FT   DNA_BIND      1    369
FT   DOMAIN      370    507
FT   DOMAIN      382    456       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      508    996       BRCT.
FT   ZN_FING      19     54       BY SIMILARITY.
FT   ZN_FING     126    164       NAD-BINDING.
FT   DOMAIN      211    214       BY SIMILARITY.
FT   DOMAIN      232    235       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT                                NUCLEAR LOCALIZATION SIGNAL 2ND PART.
SQ   SEQUENCE   996 AA;  113018 MW;  690DD36E7487298 CRC64;

Query Match             23.9%;  Score 674.5;  DB 1;  Length 996;
Best Local Similarity   33.6%;  Pred. No. 2.1e-39;
Matches  179;  Conservative 104;  Mismatches 195;  Indels  55;  Gaps  22;

Qy    25  EEDPFRSTAEALKAIPAEKRI-----IRVDPTCPLSSNPGTQVY----EDYNCTLNQTN   74

Db   490  ESKSSKSKSIYTKSVPKSWTLKIKDGLAVDPDSGLED--VAHVVSRNKEKYNVVLGITD  547

Qy    75  IENNNNKFYIIQLLQD--SNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFRE  131

Db   548  IQKNKNSFYKLQLLESDMKNRFWV-FRSWGRIGTTIGGNKLDNFSNLVDAIVQFKELYLE  606

Qy   132  KTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQ  190

Db   607  KSGNIFENRENFVKVAGRMYPIDIDYAEDS------KIDLSAEHDIKSKL-PLSV-----Q  655
```

```
Qy   191  KLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSL  250

Db   656  DIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLTEIYELIQG---GGTNA  712

Qy   251  E--ELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTV  308

Db   713  KFIDATNRFYLIPHNFGTQSPPLLDTEQVEQLRQMLDSLEIECAYSLLQTEDSKADI  772

Qy   309  EEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEG  365

Db   773  ----NPIDKHYEQLKTKLEPLDKNSEYILLQKYVNTHAETHKLYDLEVVDIFKVARQG  828

Qy   366  EEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSK  421

Db   829  EARRYKPFKKLHNRRLLWHGSRLTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADMVSK  888

Qy   422  SAGYVIGMKCGAHH--VGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP  479

Db   889  SANYC---CTSHHNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFGRGRTMPNP  943

Qy   480  TQDTELELDGQQVVVPGQPVPCPEFSSSTFSQSEVLIYQESQCRLRYLLEVH  532

Db   944  SESIIRE-DG--VEIPLGKPITNDSLKSSLL-YNEFIYDIAQVNIQYMLRMN  992


RESULT   7
PPOL_CHICK
ID   PPOL_CHICK     STANDARD;      PRT;   1011 AA.
AC   P26446;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=91340148; PubMed=1840535;
RA   Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;
RT   "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid
RT   sequence and comparison with mammalian enzyme sequences.";
RL   Gene 102:157-164(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.
RX   MEDLINE=96353841; PubMed=875499;
RA   Ruf A., Mennissier de Murcia J., de Murcia G.M., Schulz G.E.;
RT   "Structure of the catalytic fragment of poly(AD-ribose) polymerase
RT   from chicken.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO
RP   895.
RX   MEDLINE=98191351; PubMed=9521710;
RA   Ruf A., de Murcia G.M., Schulz G.E.;
RT   "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived
RT   from crystal structures and homology modeling.";
RL   Biochemistry 37:3893-3900(1998).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.
RX   MEDLINE=98239716; PubMed=9571033;
RA   Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;
RT   "The mechanism of the elongation and branching reaction of poly(ADP-
RT   ribose) polymerase as derived from crystal structures and
RT   mutagenesis.";
RL   J. Mol. Biol. 278:57-65(1998).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
```

CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC -!- SUBUNIT: HOMODIMER (Potential).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC --

DR   EMBL; X52690; CAA36917.1; -.
DR   PIR; JH0581; JH0581.
DR   PDB; 2PAW; 27-MAY-98.
DR   PDB; 1PAX; 15-MAY-97.
DR   PDB; 2PAX; 27-MAY-98.
DR   PDB; 3PAX; 27-MAY-98.
DR   PDB; 4PAX; 27-MAY-98.
DR   PDB; 1A26; 27-MAY-98.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; 3D-structure.

FT   DNA_BIND   1   370
FT   DOMAIN   371   522   AUTOMODIFICATION DOMAIN.
FT   DOMAIN   382   458   BRCT.
FT   DOMAIN   523   1011   NAD-BINDING.
FT   ZN_FING   21   56   PARP-TYPE.
FT   ZN_FING   125   162   PARP-TYPE.
FT   DOMAIN   207   209   NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN   220   225   NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES   403   403   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   404   404   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   410   410   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   411   411   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   432   432   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   434   434   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   441   441   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   442   442   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   453   453   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   454   454   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   468   468   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   481   481   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   485   485   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   488   488   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   509   509   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES   510   510   ADP-RIBOSYL[N] (POTENTIAL).

FT   MOD_RES   ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT   517   517   A -> R (IN REF. 1).
SQ   SEQUENCE   1011 AA;   113520 MW; 261AED9383139144 CRC64;

Query Match   23.9%; Score 674.5; DB 1; Length 1011;
Best Local Similarity   33.4%; Pred. No. 2.1e-39;
Matches 187; Conservative 100; Mismatches 204; Indels 69; Gaps 24;

```
Qy      7 PW---VQTEGPE-----------KKKGKQAGREDPFRSTAEALKAIPAEKRIIRVDPT   51
Db    477 PWGAEVKTEHQEVAVDGKCSKPANMKSAGKVKEEQGPSKSEKKMKLTV--KGGAAVDPD  533

Qy     52 CPLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQD-SNRFFTCWNRWGRVG    105
Db    534 SGLEDS-AHVFEKGGKIFSATLGLVDIVKGTNSYKLQLLEDDRESRYWV-FRSWGRVG   590

Qy    106 EV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQE  163
Db    591 TVIGSNKLEQMPSKEDAVEHFLNLYEEKTGNSWHSK-NFTKYPKKFYPLEIDYGQDEEAV  649

Qy    164 AVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSK  223
Db    650 RKLTVSAG---TKSKLAKP------IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSK  700

Qy    224 QQIARGFEALEALEEALKGPTDGGQSLE-ELSSHFYTVIPHNFGHSQPPPINSPELLQA   281
Db    701 RQIQSAYSILNEVQQAV---SDGGSESQILDLSNRFYTLIPHDFGMKKPPLLSNLEYIQA  757

Qy    282 KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQT  341
Db    758 KVQMLDNLLDIEVAYSLLRGGNEDGDKD---PIDINYEKLRTDIKVVDKDSEEAKIIKQ   813

Qy    342 YLEQT-GSNHRCPTLQ-HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSG   398
Db    814 YVKNTHAATHNAYDLKVVEIFRIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSOG  873

Qy    399 LRIMPH----SGGRVGKGIYFASENSKSAGYVIGMKC---GAHHVGYMFGLEVALGREHH  451
Db    874 LRIAPPEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQADPIGLILLGEVALGNMYE  928

Qy    452 INPDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQOVVVPQGQPVPCPEFSSSTFS  511
Db    929 LKNASHITKL-PKGKHSVKGLGKTAPDPTATT--TLDG--VEVPLGNGI-STGINDTCLL  982

Qy    512 QSEYLIYQESQCRLRYLLEV  531
Db    983 YNEYIVYDVAQVNLKYLLKL 1002
```

RESULT 8
PPOL_CRIGR
ID   PPOL_CRIGR   STANDARD;   PRT;   1012 AA.
AC   Q9R152;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21276334; PubMed=11382339;
RA   Ganesh A., Phillips E., Thacker J., Meuth M.;
RT   "Suppression of the radiation-sensitive phenotype of hamster irs1 and
RT   irs2 strains selected for resistance to 3-aminobenzamide.";
RL   Int. J. Radiat. Biol. 77:609-616(2001).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR

```
Matches 186; Conservative 100; Mismatches 205; Indels 59; Gaps 22:

Qy    2 APKPK--PWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNP  58
         |||    | ::|| | ::| |   ::|   |   |  :|||  | |||||
Db  493 APKGKSAAPSKKSKGLYKEEG--VNKSEKRMKLTLKGGAA-------VDPDSGLEHS-  540

Qy   59 GTQVYED---YNCTLNQTNIENNNKFYIIQLLQDSNRF-FTCWNRWRGRGEV-GQSKI  112
         |        |    :|| : |    :|| || : | :|||   |   : : |:   :|
Db  541 -AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWRGVGTVIGSNKL  599

Qy  113 NHFTRLEDAKKDFEKKFREKFKNNWAAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVDRG  171
         ||| | : | :| |  : : :|      :|| |   ||||  || : |||||||||:
Db  600 EQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGGDE--EAVKKL----  653

Qy  172 PVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFE  231
         | :    : : :|   | :|| : :| ::|  | |  ::||||||||||||||:| |
Db  654 TVKPGTK----SKLPKAVQELVGMIPDVESMKKALVEYEIDLQKMPLGKLSKRQIQAAYS  709

Qy  232 ALEALEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLAD  291
         ||  : : : : ::  ||| : ||||| |||:||  ||||:|  || |:|::| ::|||
Db  710 ILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLD  768

Qy  292 IELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNH  350
         ||:|  | : : :   |  | |||| |::| |   |: |  || ||: |  | : : |:|
Db  769 IEVAYSLLRGGSDDSSKD---PIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTH  824

Qy  351 RCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----  404
         | |      :| :::|||  ||: ::|:|:|:||||| : | | |:|| ||||   |
Db  825 NAYDLEVMDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTNFAGILSQGLRIAPPEAPV  884

Qy  405 SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDNPSLKS  461
          : ||||||||| :||  |  |||   |:     | |||:|||||| : ::    : :|
Db  885 TGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEVALGNMYELKHAS-HISK  938

Qy  462 PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES  521
          |||  : ::| ||||||: ::  ||| | :| | || : | :     || || : ::|
Db  939 LPKGKHSVKGLGKTTPDPSASITLE---GVEVPLGTGIP-SGVNDYCLLYNEYIVVDIA  993

Qy  522 QCRLRYLLEV  531
         | ||:||| :
Db  994 QVNLKYLLKL  1003


RESULT 9
PPOL_HUMAN
ID   PPOL_HUMAN     STANDARD;     PRT;    1013 AA.
AC   P09874;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+))
DE   ADP-ribosyltransferase(polymerizing) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT OR PPOL OR PARP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90091744; PubMed=2513174;
RA   Auer B., Nagl U., Herzog H., Schweiger R., Schweiger M.;
RT   "Human nuclear NAD+ ADP-ribosyltransferase(polymerizing):
RT   organization of the gene.";
RL   DNA 8:575-580(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88076933; PubMed=3120710;
RA   Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S.,
RA   Suzuki H., Nyunoya H., Miwa M., Sugimura T.;
RT   "Nucleotide sequence of a full-length cDNA for human fibroblast
RT   poly(ADP-ribose) polymerase.";
RL   Biochem. Biophys. Res. Commun. 148:617-622(1987).
```

```
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE (BY
CC   similarity).
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN (BY
CC   SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (Potential).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF168781; AAD45817.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET        0      0
FT   DNA_BIND        1    371       BY SIMILARITY.
FT   DOMAIN        372    522       BY SIMILARITY.
FT   DOMAIN        386    475       AUTOMODIFICATION DOMAIN (BY SIMILARITY).
FT   DOMAIN        523   1012       NAD-BINDING (BY SIMILARITY).
FT   ZN_FING        20     55       PARP-TYPE.
FT   ZN_FING       124    161       PARP-TYPE.
FT   DOMAIN        206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                  (POTENTIAL).
FT   DOMAIN        220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                  (POTENTIAL).
FT   MOD_RES       406    406       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       412    412       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       443    443       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       455    455       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       470    470       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       483    483       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       487    487       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       490    490       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       511    511       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES       518    518       ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE  1012 AA;  112400 MW;  8D617C4DBF0CB0F7 CRC64;

Query Match            23.9%;
Best Local Similarity  33.8%;
Score 673.5;  DB 1;  Length 1012;
Pred. No. 2.5e-39;
```

```
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88058958; PubMed=2824474;
RA   Kurosaki T., Ushiro H., Mitsuuchi Y., Suzuki S., Matsuda M.,
RA   Matsuda Y., Katunuma N., Kangawa K., Matsuo H., Hirose T.,
RA   Inayama S., Shizuta Y.;
RT   "Primary structure of human poly(ADP-ribose) synthetase as deduced
RT   from cDNA sequence.";
RL   J. Biol. Chem. 262:15990-15997(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88068596; PubMed=2891139;
RA   Cherney B.W., McBride O.W., Chen D., Alkhatib H., Bhatia K.,
RA   Hensley P., Smulson M.E.;
RT   "cDNA sequence, protein structure, and chromosomal location of the
RT   human gene for poly(ADP-ribose) polymerase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:8370-8374(1987).
RN   [5]
RP   SEQUENCE OF 440-1013 FROM N.A.
RX   MEDLINE=87298455; PubMed=3113420;
RA   Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA   Miwa M.;
RT   "Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and
RT   expression of its gene during HL-60 cell differentiation.";
RL   Biochem. Biophys. Res. Commun. 146:403-409(1987).
RN   [6]
RP   ERRATUM.
RA   Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA   Miwa M.;
RL   Biochem. Biophys. Res. Commun. 148:1549-1550(1987).
RN   [7]
RP   SEQUENCE OF 1-94 FROM N.A.
RX   MEDLINE=91099327; PubMed=2125269;
RA   Yokoyama Y., Kawamoto T., Mitsuuchi Y., Kurosaki T., Toda K.,
RA   Ushiro H., Terashima M., Sumimoto H., Kuribayashi I., Yamamoto Y.,
RA   Maeda T., Ikeda H., Sagara Y., Shizuta Y.;
RT   "Human poly(ADP-ribose) polymerase gene. Cloning of the promoter
RT   region.";
RL   Eur. J. Biochem. 194:521-526(1990).
RN   [8]
RP   SEQUENCE OF 1-39 FROM N.A.
RX   MEDLINE=90211250; PubMed=2108670;
RA   Ogura T., Nyunoya H., Takahashi-Masutani M., Miwa M., Sugimura T.,
RA   Esumi H.;
RT   "Characterization of a putative promoter region of the human
RT   poly(ADP-ribose) polymerase gene: structural similarity to that of
RT   the DNA polymerase beta gene.";
RL   Biochem. Biophys. Res. Commun. 167:701-710(1990).
RN   [9]
RP   SEQUENCE OF 1-39 FROM N.A.
RA   Herzog H., Schneider R., Hirsch-Kauffmann M., Schnitzer D.,
RA   Schweiger M.;
RL   Submitted (JUL-1991) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=90222155; PubMed=2109322;
RA   Gradwohl G., Menissier de Murcia J., Molinete M., Simonin F.,
RA   Koken M.H.M., Hoeijmakers J.H.J., de Murcia G.M.;
RT   "The second zinc-finger domain of poly(ADP-ribose) polymerase
RT   determines specificity for single-stranded breaks in DNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:2990-2994(1990).
RN   [11]
RP   ANALYSIS OF ZINC FINGERS.
RX   MEDLINE=91072398; PubMed=2123876;
RA   Ikejima M., Noguchi S., Yamashita R., Ogura T., Sugimura T.,
RA   Gill D.M., Miwa M.;
RT   "The zinc fingers of human poly(ADP-ribose) polymerase are
RT   differentially required for the recognition of DNA breaks and nicks
RT   and the consequent enzyme activation. Other structures recognize
RT   intact DNA.";
RL   J. Biol. Chem. 265:21907-21913(1990).
RN   [12]

RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=9103546; PubMed-2121735;
RA   Simonin F., Menissier de Murcia J., Poch O., Muller S., Gradwohl G.,
RA   Molinete M., Penning C., Keith G., de Murcia G.M.;
RT   "Expression and site-directed mutagenesis of the catalytic domain of
RT   human poly(ADP-ribose)polymerase in Escherichia coli. Lysine 893 is
RT   critical for activity.";
RL   J. Biol. Chem. 265:19249-19255(1990).
RN   [13]
RP   NUCLEAR LOCALIZATION SIGNAL.
RX   MEDLINE=92371433; PubMed=1505517;
RA   Schreiber V., Molinete M., Boeuf H., de Murcia G.M.,
RA   Menissier de Murcia J.;
RT   "The human poly(ADP-ribose) polymerase nuclear localization signal is
RT   a bipartite element functionally separate from DNA binding and
RT   catalytic activity.";
RL   EMBO J. 11:3263-3269(1992).
RN   [14]
RP   MUTAGENESIS OF CATALYTIC DOMAIN.
RX   MEDLINE=97461532; PubMed=9315851;
RA   Rolli V., O'Farrell M., Menissier-de Murcia J., de Murcia G.M.;
RT   "Random mutagenesis of the poly(ADP-ribose) polymerase catalytic
RT   domain reveals amino acids involved in polymer branching.";
RL   Biochemistry 36:12147-12154(1997).
-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
-!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
-!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
-!- SUBUNIT: HOMODIMER (Potential).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
-!- SIMILARITY: BELONGS TO THE PARP FAMILY.
-!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
--------------------------------------
```

```
DR   EMBL; X16674; CAA34663.1; -.
DR   EMBL; M18112; AAA60137.1; -.
DR   EMBL; J03473; AAB59447.1; -.
DR   EMBL; M17081; AAA51599.1; ALT_SEQ.
DR   EMBL; M32721; AAA60155.1; -.
DR   EMBL; M29786; AAA51663.1; -.
DR   EMBL; M29545; AAA51663.1; JOINED.
DR   EMBL; M29766; AAA51663.1; JOINED.
DR   EMBL; M29767; AAA51663.1; JOINED.
DR   EMBL; M29768; AAA51663.1; JOINED.
DR   EMBL; M29769; AAA51663.1; JOINED.
DR   EMBL; M29770; AAA51663.1; JOINED.
DR   EMBL; M29771; AAA51663.1; JOINED.
DR   EMBL; M29772; AAA51663.1; JOINED.
DR   EMBL; M29773; AAA51663.1; JOINED.
DR   EMBL; M29774; AAA51663.1; JOINED.
DR   EMBL; M29775; AAA51663.1; JOINED.
DR   EMBL; M29776; AAA51663.1; JOINED.
DR   EMBL; M29777; AAA51663.1; JOINED.
DR   EMBL; M29778; AAA51663.1; JOINED.
DR   EMBL; M29779; AAA51663.1; JOINED.
DR   EMBL; M29780; AAA51663.1; JOINED.
```

```
DR   EMBL; M29781; AAA51663.1; JOINED.
DR   EMBL; M29783; AAA51663.1; JOINED.
DR   EMBL; M29784; AAA51663.1; JOINED.
DR   EMBL; M29785; AAA51663.1; JOINED.
DR   EMBL; M29544; AAA51663.1; JOINED.
DR   EMBL; M29782; AAA51663.1; JOINED.
DR   EMBL; X56140; CAA39606.1; -.
DR   EMBL; X56141; CAA39606.1; JOINED.
DR   EMBL; M60436; AAA60000.1; -.
DR   PIR; A26901; A26901.
DR   PIR; A28498; A28498.
DR   PIR; A29725; A29725.
DR   PIR; A35635; A35635.
DR   PIR; A33321; A33321.
DR   PIR; B33321; B33321.
DR   PIR; A39976; A39976.
DR   PIR; S14010; S14010.
DR   HSSP; P26446; 1A26.
DR   Aarhus/Ghent-2DPAGE; 1620; NEPHGE.
DR   MIM; 173870; -.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.

Query Match             23.6%;  Score 666.5;  DB 1;  Length 1013;
Best Local Similarity   33.4%;  Pred. No. 7.7e-39;
Matches  186;  Conservative  101;  Mismatches  207;  Indels  63;  Gaps  24;

Qy      7 PW--------VQTEGPEKKKGRQAGR------EEDPFRSTAEALKAIPAEKRIIRVDPTC 52
                    : :         ||               ||::  |    ::   |||||||
Db    479 PWGAEVKAEPVEVVAPRGKSGAALSKKSKGQVKEEGINKSEKRMKL--TLKGGAAVDPDS 536

Qy     53 PLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWGRVGE 106
            |     |         :: |||     :||:|| |   |:|::  :   |||||
Db    537 GLEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYW-IFRSWGRVGT 593

Qy    107 V-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEA 164
          | |::|     |||  ||::|:: || + ||:||| |||||:| |:   ::  ||:| ||
Db    594 VIGSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EA 650

Qy    165 VVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQ 224
          : | ::|   |: :   ::|: |||| ||:  :::::| :|:|| : |::|||||:|||
Db    651 VKKLITVNP-GTKSKLPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKR 703

Qy    225 QIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKD 284
          || : : :  ::    :      ::   |:|   :| |:::| |::   :::|
Db    704 QIQAAYSILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVE 762

Qy    285 MLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLE 344
          ||:: :  :|  :  |:|||       |:  | |  :|  :   |   || ::: : |:
Db    763 MLDNLLDIEVAYSLRGGSDDSSKD---PIDVNYEKLKTDIKVVDRDSEEAEIIRKYVK 818

Qy    345 QT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRI 401
          |  | |||  ||   |:||::: ::|| |  ||  |::|:|| ||  |||||:|||||
Db    819 NTHATTHNAYDLEVIDFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRI 878

Qy    402 MPH----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINT 454
             :     |||||| ||||||      ||  ||:||:          :| ||||||| :
Db    879 APPEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEVALGNMYELKH 933

Qy    455 DNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSE 514
          ||  | | |||||  :  |  :| :  | ||  |:  |  |   | | |  ::| |:
Db    934 AS-HISKLPRGKHSVKGLGKTTPDPS--ANISLDG--VDVPLGTGI-SSGVNDTSLLYNE 987

Qy    515 YLIYQESQCRLRYLLEV 531
          ||: :       ||| :
Db    988 YIVYDIAQVNLKYLLKL 1004
```

RESULT   10

PPOL_RAT
```
ID   PPOL_RAT       STANDARD;       PRT;       1013 AA.
AC   P27008; O35937;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Monocytes;
RX   MEDLINE=98046546; PubMed=9385436;
RA   Beneke S., Meyer R., Buerkle A.;
RT   "Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly
RT   (ADP-ribose) polymerase.";
RL   Biochem. Mol. Biol. Int. 43:755-761(1997).
RN   [2]
RP   REVISION TO 811.
RA   Beneke S., Meyer R., Buerkle A.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-11 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=92290013; PubMed=1601134;
RA   Potvin F., Thibodeau J., Kirkland J.B., Dandenault B.,
RA   Duchaine C., Poirier G.G.;
RT   "Structural analysis of the putative regulatory region of the rat
RT   gene encoding the poly(ADP-ribose) polymerase)
RL   FEBS Lett. 302:269-273(1992).
RN   [4]
RP   SEQUENCE OF 514-1013 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=90027702; PubMed=2508731;
RA   Thibodeau J., Gradwohl G., Dumas C., Clairoux-Moreau S., Brunet G.;
RT   "Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase
RT   catalytic domain and analysis of mRNA levels during the cell cycle.";
RL   Biochem. Cell Biol. 67:653-660(1989).
-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
-!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
-!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
-!- SIMILARITY: BELONGS TO THE PARP FAMILY.
-!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```
```
DR   EMBL; U94340; AAC53544.1; -.
DR   EMBL; X65496; CAA46477.1; -.
DR   EMBL; X65497; CAA46478.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
```

```
Db   821  HATTHNAYDLEVIDIFKIEREGESQRYKFPRQLHNRRLLMHGSRTNFAGILSQGLRIAP 880
                :    |||||||| | |  ||| ::|
Qy   404  H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDN 456
                :  ||||||||                 :  |||||||
Db   881  PEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEVALGNMYELKHAS 935
                :  | || ||| ||| ::     : || | :  ||| ||
Qy   457  PSLKSPPPGPDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSSTFSQSEYL 516
                :   | || ||| :|||:  ::|| | :  :||:  :|:
Db   936  -HISKLPKGKHSVKGLGKTAPDPS--ASITLDG--VEVPLGTGIP-SGVNDTCLLYNEYI 989
                : | ||||  :|:|::  ::
Qy   517  IYQESQCRLRYLLEV 531
                :| ::|    |:|::
Db   990  VYDIAQVNLKYLLKL 1004
```

```
RESULT 11
PPOL_XENLA
ID   PPOL_XENLA        STANDARD;      PRT;    998 AA.
AC   P31669;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Saulier-Le Drean B.M.;
RL   Thesis (1992), University of Rennes, France.
RN   [2]
RP   SEQUENCE OF 742-876 FROM N.A.
RX   MEDLINE=9327538; PubMed=8503897;
RA   Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,
RA   Miwa M.;
RT   "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)
RT   polymerase from Xenopus laevis and cherry salmon using heterologous
RT   oligonucleotide consensus sequences.";
RL   Biochem. Biophys. Res. Commun. 193:119-125(1993).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,
CC       AND BRAIN. LOW IN LIVER.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z12139; CAA78126.1; -.
DR   EMBL; D13810; BAA02966.1; -.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   DNA_BIND      1    372.
FT   DOMAIN      385    461       BRCT.
FT   DOMAIN      373    523       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      524   1013       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       PARP-TYPE.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   DOMAIN        1      1       BLOCKED (BY SIMILARITY).
FT   MOD_RES     407    407       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     413    413       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     435    435       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     437    437       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     456    456       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     484    484       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     491    491       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     513    513       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     519    519       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    638    638       Y -> H (IN REF. 4).
FT   CONFLICT    641    641       E -> A (IN REF. 4).
FT   CONFLICT    752    752       N -> D (IN REF. 4).
SQ   SEQUENCE   1013 AA;  112529 MW;  AA566F2B29BE97C0 CRC64;
```

```
Query Match             23.6%;  Score 665.5;  DB 1;  Length 1013;
Best Local Similarity   33.5%;  Pred. No. 9.1e-39;
Matches 186;  Conservative 99;  Mismatches 203;  Indels 67;  Gaps 23;

Qy     1  MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRI------IRVDPTCP 53
             |:    |   :|  : |       |       ||| | :|            |||| |
Db   493  VVPKGK----SAAPSKKSKGAVKEEGVNKS--------EKRMKLTLKGGAAVDPDSG 537
             : | ||                :|:                 |: :|||||
Qy    54  LSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEV- 107
            | |  |        :  :||  :| :| ||||| : |  :|: |  |||  ||||||
Db   538  LEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLESDKESRYWIFRSWGRVGTVI 595
             | ||  | || ||| : |    ||     :| : ||| :|  |   |||   ||||
Qy   108  GQSKINHFTRLEDAKKDFEKKFREKTKNNWAAERDHFVSHPGKYTLIEVQ-AEDEAQEAVV 166
             | ||  |  ||  | :|| | :  |  |   :  | : ||    |:|   ||
Db   596  GSNKLEQMPSKDAVEHFMKLYEEKTGNAWHSK-NFTKYPKFYPLEIDYGGDE--EAVK 652
             :|:|                :|   | | |:  | ||| ||   ||   :| |  | |
Qy   167  KVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPIGKLSKQQI 226
            |:                 |:|  | |  | ||| |  | :|:|   | :|  |  ||
Db   653  KL---AVKPGTK---SKLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQI 705
             |:   |  |:|       :| | :|  |:  |:  | |:| |  |||  |||| |||
Qy   227  ARGFEALEALEALGKPTDGGQSGSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDML 286
             |  |  | | | | : :: ||    |  |  |  ||  :| | :|: ||  | :| ::||
Db   706  QAAYSILSEVQQAVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNTDSVQAKVEML 764
             :  || | ||   | :|| |   ::|  | ||| || | ::| | ||   :| ||:|:||
Qy   287  LVLADIELAQALQAVSEQEKIVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 346
            | | |  || |    |:    ::: |       |:|  | |::  |  ||:: |  :| |
Db   765  DNLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNT 820
             ||:|||  :|| || | : :         |: | ||:| ||  |:| :| ||  |: |||
Qy   347  -GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKIGNRKLLWHGTNMAVVAAILFSGLRIMP 403
             |:   | |     |:: |||  | | |   | : | ||   :| ||||||:  |||:||
```

```
DR   PIR; S31735; S31735.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   NON_TER       1      1
FT   DNA_BIND     <1    356
FT   DOMAIN      357    507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      369    445       BRCT.
FT   DOMAIN      508    998       NAD-BINDING.
FT   ZN_FING       8     43       PARP-TYPE.
FT   ZN_FING     111    148       PARP-TYPE.
FT   DOMAIN      193    195       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      207    212       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     391    391       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     397    397       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     419    419       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     428    428       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     429    429       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     447    447       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     454    454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     467    467       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     471    471       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     477    477       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     495    495       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     496    496       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     503    503       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    746    746       Q -> E (IN REF. 2).
SQ   SEQUENCE    998 AA;  111126 MW;  F5A25E4A3366BAE7 CRC64;

Query Match             23.1%;  Score 653;  Pred. No. 6.6e-38;  DB 1;  Length 998;
Best Local Similarity   32.5%;
Matches  177;  Conservative 102;  Mismatches 214;  Indels 52;  Gaps 20;

Qy      3  PKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQV   62
           |  |   |  :| |:  |:  |   ||  :  |   || : :| : |  |
Db    480  PSSGPVAGKSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDSEL---EDSCHVLETGG---  533

Qy     63  YEDYNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRKWGRVGEV-GQSKINHFTRLED  120
           :  |: |  ||   :| |   ||   | :  ||  |:|   |      |  | : :|
Db    534  -KIFSATLGLVDITRGTNSYYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKED  592

Qy    121  AKDFEKKFREKTKNNWAEERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV  180
           ::||   | : :  | ||    | : ||:|| : :|  |   |
Db    593  AIEHFLNLYQDKTGNAW-HSPNFTKYPKKFYPLEIDYGGE-EDVVKKLSVG-AGTKSKLA  649

Qy    181  QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL  240
           :|  | || |   ::|||  | |:  ||: :|:| || |:|||  |   |:
Db    650  KP-----VQELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAV  703

Qy    241  KGPTDGGQSLEE----LSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELA  295
           | |  | :  |     | | ||: :      |  ||:| :| : ::|  ||::| | |
Db    704  ------SESLSEARLLDLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVA  757

Qy    296  QAL--QAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRC  352
           :|   || :| |       :|:   |  | |:| || :| :  || :  |:|  |  |
Db    758  YSLLRGGADDGEK-----DPIDVKYEKIKTDIKVVAKDSEESRIICDVVKNTHADTHNA  811

Qy    353  PTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SG  406
```

```
Db    812  YDLEVLEIFKIDREGEYQRYKPFKQLHRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTG  871
           |:  |:::::|||    :  | :|||||   ||:||: |   |   : |: | :
Qy    407  GRVGKGIYFASENSKSAGYVIGMKCGAHHVGMFLGEVALGREHHINTDNPSLKSPPPGF  466
           |||||| :|||  :  | | || |:  : |:|||||| : : :: : |  :  |
Db    872  YMFGKGIYFADMVSKSANYCHAMP--GSPIGLILLGEVALGNMHELKAASQITKL-PKGK  928
           |||||| :| | : |  | |   |   |  |||| ||||: :   |:| :|  :  |||
Qy    467  DSVIARGHTEPDTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLR  526
           |   | | | |  :| |||:|| | : | :|  :| : | : || :  |   :|  |
Db    929  HSVKGLGRTAPDFS--ATVQLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLK  983

Qy    527  YLLEV  531
           |||::
Db    984  YLLKL  988


RESULT 12
PPOL_MOUSE
ID   PPOL_MOUSE      STANDARD;      PRT;   1012 AA.
AC   P11103; Q9JLX4; Q9QV03;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1) (msPARP).
GN   ADPRT OR ADPRT1 OR ADPRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=BXSB;
RX   MEDLINE=89263780; PubMed=2498841;
RA   Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;
RT   "Sequence and organization of the mouse poly (ADP-ribose) polymerase
RT   gene.";
RL   Nucleic Acids Res. 17:3387-3401(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC   STRAIN=129/Sv X C57BL/6; TISSUE=Fibroblast;
RX   MEDLINE=20270268; PubMed=10809783;
RA   Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;
RT   "Characterization of sPARP-1. An alternative product of PARP-1 gene
RT   with poly(ADP-ribose) polymerase activity independent of DNA strand
RT   breaks.";
RL   J. Biol. Chem. 275:15504-15511(2000).
RN   [3]
RP   KNOCK-OUT.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jaeger S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "on the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC       short form/sPARP; may be produced by alternative initiation.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
```

```
Query Match                       23.0%;    Score 649.5;  DB 1;    Length 1012;
Best Local Similarity      32.4%;   Pred. No. 1.2e-37;
Matches 180;  Conservative 103;  Mismatches 205;  Indels 67;  Gaps 22;

Qy    1 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRI------IRVDPTCP 53
          :||:|                                          |||
Db  492 VAPRGK-----SAAPSKKSKGCFKEEGVNKS----------EKRMKLTLKGGAAVDPDSG 536

Qy   54 LSSNPGTQYED---YNCTLNQTNIENNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV- 107
          :||:|          |:  :|    : |||: |||:|    :|||:|     |:|
Db  537 LEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRLGTVI 594

Qy  108 GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVV 166
          |  :|:            ::  :|: ||| |    :|: :| :  |||   |:   |||
Db  595 GSNKLEQMPSKEEAVEQFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVK 651

Qy  167 KVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQI 226
          |: |    |   :  :         :|| || |     : | |      |:  |||:|:||
Db  652 KLIVKP-GTKSKLPKP------VQELVGMIFDVDSMKKALVEYEIDLQKMPLGKLSRRQI 704

Qy  227 ARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDML 286
          |:  :  : |:  | : | |   | ::| |     |:|  : :|:||: :|||: |||  |
Db  705 QAAYSILSEVQQPVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEML 763

Qy  287 LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 346
          |:||  |  :| || |: | :   :|     | ||::|  ::|||   :|   |:|| |||
Db  764 DNLLDIEVAYSLLRGGSDDSSKD---PIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNT 819

Qy  347 -GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGNMAVVAAILTSGLRIMP 403
           ||:| |       |||::: : ||: | ::|||: ||||||| ||  || : |||:|
Db  820 HATTHNAYDLEVIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSOGLRIAP 879

Qy  404 H------SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDN 456
          |          |||||||||:|:||:||||||||| |||      |||||| |||:||||||||
Db  880 PEAPVTGYMFGKGIYFADMVSKSANY------CHTSQGDPIGLIMLGEVALGNMYELKHAS 934

Qy  457 PSLKSPPPGFDSVIARGHTEPDPTQDTELEDGGQVVPVPQGQPVPCPEFSSSTFSQSEYL 516
          ||| ||| | ||:|  |    |||  |:::  |||   |  |  | ||| :|:: |:|:|:
Db  935 -HISKLPKGKHSVKGLGKTTPDPSASITLE----GVEVPLGTGIP-SGVNDTALLYNEYI 988

Qy  517 IYQESQCRLRYLLEV 531
          :|:|||  |:||:|
Db  989 VYDIAQVNLKYLLKL 1003
```

```
RESULT 13
PPOL_DROME
ID   PPOL_DROME      STANDARD;       PRT;   994 AA.
AC   P35875; Q9W5Q5; Q9W5S1;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   PARP OR CG17696/CG17718.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93234521; PubMed=8475096;
RA   Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,
RA   Sugimura T., Miwa M.;
RT   "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:
RT   leucine zipper in the auto-modification domain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND
RP   TISSUE SPECIFICITY.
RC   STRAIN=CANTON-S;
RX   MEDLINE=98234380; PubMed=9555614;
```

```
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```

```
DR   EMBL; X14206; CAA32421.1; -.
DR   EMBL; AF126717; AAF61293.1; ALT_INIT.
DR   PIR; S04200; S04200.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1340806; Adprtl.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative initiation.
FT   INIT_MET      0      0
FT   CHAIN         1   1012       POLY [ADP-RIBOSE] POLYMERASE-1, LONG
FT                                ISOFORM.
FT   CHAIN       521   1012       POLY [ADP-RIBOSE] POLYMERASE-1, SHORT
FT                                ISOFORM.
FT   INIT_MET    521    521       FOR SHORT ISOFORM.
FT   DNA_BIND    371    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      372    460       BRCT.
FT   DOMAIN      384   1012       NAD-BINDING.
FT   DOMAIN      523   1012       PARP-TYPE.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   DOMAIN      220    225       BY SIMILARITY.
FT   MOD_RES     406    406       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     412    412       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     443    443       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     447    447       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     455    455       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     483    483       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     487    487       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     490    490       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     511    511       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     518    518       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    590    590       L -> V (IN REF. 2).
FT   CONFLICT    607    607       E -> D (IN REF. 2).
FT   CONFLICT    611    611       Q -> H (IN REF. 2).
FT   CONFLICT    628    628       N -> D (IN REF. 2).
FT   CONFLICT    678    678       D -> E (IN REF. 2).
FT   CONFLICT    702    702       R -> E (IN REF. 3).
FT   CONFLICT    716    716       Q -> E (IN REF. 3).
FT   CONFLICT    757    757       Q -> L (IN REF. 2).
FT   CONFLICT    856    856       R -> F (IN REF. 2).
FT   CONFLICT    981    981       A -> C (IN REF. 2).
SQ   SEQUENCE  1012 AA;  112968 MW;  4354C3E5F01B9439 CRC64;
```

```
RA   Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;
RT   "Genomic organization of Drosophila poly(ADP-ribose) polymerase and
RT   distribution of its mRNA during development.";
RL   J. Biol. Chem. 273:11881-11886(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES
CC       OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM
CC       IN LATER EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN
CC       EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
DR   EMBL; D13806; BAA02964.1; -.
DR   EMBL; AF051548; AAC24518.1; -.
DR   EMBL; AF051544; AAC24518.1; JOINED.
DR   EMBL; AF051545; AAC24518.1; JOINED.
DR   EMBL; AF051546; AAC24518.1; JOINED.
DR   EMBL; AF051547; AAC24518.1; JOINED.
DR   EMBL; AE002935; AAF45400.1; -.
DR   EMBL; AE002666; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE002892; AAF45545.2; ALT_SEQ.
DR   PIR; A47474; A47474.
DR   HSSP; P26446; 1A26.
DR   FlyBase; FBgn0010247; Parp.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.
FT   DNA_BIND       1      367
FT   DOMAIN       368      507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN       380      454       BRCT.
FT   DOMAIN       508      994       NAD-BINDING.
FT   ZN_FING       19       54       PARP-TYPE.
FT   ZN_FING      123      161       PARP-TYPE.
FT   DOMAIN       208      210       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN       223      228       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   VARSPLIC     376      564       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   994 AA;   113791 MW;   ACA85A270DD29E08 CRC64;

Query Match                       22.6%;
Best Local Similarity   32.2%;         Score 638;  DB 1;  Length 994;  Pred. No. 7.4e-37;
Matches 168;  Conservative 100;  Mismatches 195;  Indels 58;  Gaps 19;

Qy    37 KAIPAEKRI-----IRVDPTCPLSSNPGTQVYED----YNCTLNQTNIENNNKFYIIQL 87
          |::|        :|||     : :|: |  :  :::   :  |:|::|  | ::: :||
Db   500 KSMPVSRTFKVKDGLAVDPDSGLEDI--AHYVDSNNKYSVVLGLTDIQRNKNSYYKVQL 557

Qy    88 LQ-DSNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVS 145
          || |:::| |||:|    |::|:|:   ||:::|| ::|::| ::||
Db   558 LKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVK 617

Qy   146 HPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNT 205
          |: |:| :|: :||    : :|| :     ||
Db   618 RTGRMYPIEIQYDDD--QKLVKHE------SHFFTSKLEISVQNLIKLIFDIDSMNKT 667

Qy   206 MALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNF 265
          :  :| ||   |: | | :  |::| | ||:  :   |     |::||:|||
Db   668 LMEFHIDMDKMPLGKLSAHQIQSAYRVKEIYNVLECGSNTAK-LIDATNRFYTLIPHNF 726

Qy   266 GHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQ 325
          |  | | :| :||| |:  :||     ||| ||   |:     :  |:|| || :|::|
Db   727 GVQLPLIETHQQIEDLRQMLDSLAEIEVAYSI----IKSEDVSDACNPLDNHYAQIKTQ 782

Qy   326 LQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLL 382
          |:   || |  :||   || : :|||:  || |   :|:||   || |||||||:|||
Db   783 LVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARRFKPFKKLHNRKLL 842
```

```
Qy    253  LSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVL-----ADIELAQALQAVSEQEKT  307
           ::.:::::::: .:::::: :::::::::::
Db    239  INK-YYSLIPHSFGCVPPKIDSHAKIQAERELLDALKGSIEASLELKDLKKTASSKD--  295

Qy    308  VEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQ--HIWKVN-  362
           :: :: .:  :: : : .:::: :::  :.:
Db    296  ------IYQRLYERLPCHLEPVSE------------EIAGKIGDCLAMRGPTHCYKLSLI  337

Qy    363  ------------------QEGEEDRFQAHSKLG--------NRKLLWHGTNMAVVAAIL  395
                             .::::::  :           ::::: :::: :::
Db    338  DAFELKDPNEIPTEAPVEVQEVPKKRGKKSTKTAAPTVPPTTKRLLWHGTRVTNVFSIL  397

Qy    396  TSGLR--IMPHSGRVGKGIYFASENSKSAGYVIGMKC-GAHHVGYMFLGEVALGR----  448
           :::    :::: :: :::::::::::::::  :::::  :::: ::  ::.
Db    398  MNGLQFPVGDRCGLMFGNGVYFANVPTKSSANYC----CPEASKRVFMLLCEVETANPLVL  453

Qy    449  -EHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSQ  507
            :: .: ::   : :::: ::::: :::::::::::::: :   :: .:::
Db    454  YESEIDADEKMEKAKK---TSVYAAGKHTP--RDT-VEING-------------IPAFKSNL  496

Qy    508  STFSQ----------SEYLIYQESQCRLRYLLEV  531
           : :           :::::::: ::::::::::::
Db    497  ETIEEETRLLYDEYVMFNKEHFKIKYVVEV  526
```

```
RESULT  15
PPOV_HUMAN
ID   PPOV_HUMAN        STANDARD;      PRT;   1724 AA.
AC   Q9UKK3; O75903; Q9H1M6; Q14682;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Vault poly(ADP-ribose) polymerase (EC 2.4.2.30) (VPARP) (193-kDa vault
DE   protein) (PARP-related/alphai-related H5/proline-rich) (PH5P).
GN   ADPRTL1 OR PARPL OR KIAA0177.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 306-319.
RX   MEDLINE=99408776; PubMed=10477748;
RA   Kickhoefer V.A., Siva A.C., Kedersha N.L., Inman E.M., Ruland C.,
RA   Streuli M., Rome L.H.;
RT   "The 193 kDa vault protein, VPARP, is a novel poly(ADP-ribose)
RT   polymerase.";
RL   J. Cell Biol. 146:917-928(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=20112770; PubMed=10644454;
RA   Still I.H., Vince P., Cowell J.K.;
RT   "Identification of a novel gene (ADPRTL1) encoding a potential
RT   poly(ADP-ribosyl)transferase protein.";
RL   Genomics 62:533-536(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Tromans A.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 94-1724 FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=96281124; PubMed=8724849;
RA   Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. V.
RT   The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 3:17-24(1996).
RN   [5]
RP   DISCUSSION OF SEQUENCE.
RX   MEDLINE=99198702; PubMed=10100603;
RA   Jean L., Risler J.-L., Nagase T., Coulouarn C., Nomura N.,
RA   Salier J.-P.;
```

```
Qy    383  WHGTNMAVVAAILTSGLRI----MPHSGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY  438
           :::: ::.:   : ::::      ::::: ::::::::::::::::: :::::: :
Db    843  WHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFADMVSKSANYCCTSQ--QNSTGL  900

Qy    439  MFLGEVALG------REHHIN--TDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQ  490
           : :::::        ::.::    ::: :::::: : ::::: :   :. ::  ::
Db    901  MLLSEVALGDMMECTSAKYINKLSNNK-------HSCFGRGRTMPDPTK-SYIRSDG-  949

Qy    491  QVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEV  531
           : ::  :: : ::::  :::::::::::::: ::::::::::
Db    950  -VEIPYGETI-TDEHLKSSLLYNEYIVYDVAQVNIQYLFRM  988
```

```
RESULT  14
YQN4_CAEEL
ID   YQN4_CAEEL        STANDARD;      PRT;    538 AA.
AC   Q09525;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.
GN   E02H1.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Smith A.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-
CC       RIBOSYLTRANSFERASE (EC 2.4.2.30).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z47075; CAA87379.1; -.
DR   HSSP; P26446; 1A26.
DR   WormPep; E02H1.4; CE01539.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Hypothetical protein.
SQ   SEQUENCE  538 AA;  61268 MW;  3144E25465FC7341 CRC64;
```

```
Query Match        13.9%;  Score 393.5;  DB 1;  Length 538;
Best Local Similarity  26.1%;  Pred. No. 3.8e-20;
Matches  149;  Conservative  90;  Mismatches  172;  Indels  159;  Gaps  27;

Qy     66  YNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDF  125
           :..: :::: :::::::  :::  :::   ::. : ::: :  ::::::: .:::
Db     12  YKVHLCKTNIAQNNNKFYDMELLDEGGDFIVKLIN-GRIGYRGVTQLKDFDDLDRAKKFF   70

Qy    126  EKKFREKTKNNWAERDHFVSHPGKYTLIEV--------QAEDEAQEAVVKVD-------  169
           :::::::::   :::  ::  :::::: :          :::::::: :::
Db     71  ESKFYERTHLHWEERDD-EPVPNKYAYVELATNARQTEKEVKKEEPEPEPKVDEKNTRGR  129

Qy    170  -RGPV--RTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDV-----KKM----  216
            :::    :   :   :     :: ::   : :  :  :  : ::         :::
Db    130  KKRGIVKEKKEIKKEEP--VEEVNEKL-------KELMK---CICDEDVHLGLLKQLKFN  178

Qy    217  -----PLGKLSKQQIARGFEALEALEEALKGPT-----------------DGGQSLEE  252
                :: : ::  ::  : :  : :                       . ::::::
Db    179  EAFGRPIDCLSLAQLTTGYEILSKIEESIGGKSARRSTRGRPRVADRVLAVKSDGPSLHD  238
```

Query Match   9.1%; Score 258; DB 1; Length 1724;
Best Local Similarity   22.8%; Pred. No. 5.5e-10;
Matches 127; Conservative   92; Mismatches 193; Indels 146; Gaps 25;

```
Qy    42 EKRIRV--DPTCPLSSNP------GTQVYEDYNCTLNQTNIEN------NNNKFYI  84
            ||||::|  ||:|| |.|      ::|:|| :|.:| :: :|:|      |
Db    83 EKRLLDVKNYDPYKPLDITPPPDQKASSSEVKTEGLCPDSATEEDTVELTEFGMQNVEI 142

Qy    85 IQLLQDSNRFFTCWNRWGRVG-EVGQSKI------------------NHF------TR-  117
            ||| |     :| : ||| | :       |  : :||  ||           :||
Db   143 PHLPQDFE--VAKYNTLEKVGMEGCQEAVVVELQCSRDSRDCPFLISSHFLLDDGMETRR 200

Qy   118 ------LEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDR 170
                  ||:| ||| |:| || ::|  |:     :|:
Db   201 QFAIKKTSEDASEYFENYIEELKKQGFLLREHF------TPEATQLASEQLQALL----- 249

Qy   171 GPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLD-VKKMPLGKLSKQQIARG 229
            :|:  : :| |: ::||  ::|:
Db   250 -----LEEVMNSSTLSQEVSDLVEMIWAE-----ALGHLEHMLLKPVNRISLNDVSKA 297

Qy   230 FEALEALEEALK-GPTDGGQSLEELSSHFYTVIPHNFGHSOQPPINSPELLQAKKDMLLV 288
            |  :| |||||  ||||  : :|||||::| | :|| || || :| |:||
Db   298 EGILLLVKAALKNGET--AEQLQKMTEFYRLIPHK--GTMPKEVNL-GLLAKKADLCQL 352

Qy   289 LADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEY-KVIQTYLEQTG 347
            : :|:        : :   |::|  |  :|: ||:    | :||  |: |:
Db   353 IRDM-------VNVCETNLSKPNPPSLAKYRALRCKIEHVEQNTEEFLRVRKEVLQNHH 404

Qy   348 SNHRCPTLOHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGL--RIMPH 404
            |:|  ::|:    |:  ||||  |: :    :|||  ||  ::    :  ::
Db   405 SKSPVDVLQ-IFRVGRVNETTEFL--SKLGNVRPLLHGSPVQNIVGILCRGLLLPKVED 461

Qy   405 SG------GRVGKGIYFASENSKSAGYVIGMKCGAHHVG-----YMFLGEVALGREHHI 452
            |       |:||| ||::||: |:|| ||   | |:       |  |||:|   |  :
Db   462 RGVQRTDVGNLGSGIYFSDSLSTSIKY-------SHPGETDGTRLLLICDVALGKCMDL 513

Qy   453 NTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQ 512
            : :|: || || :|:  :: || ::||:
Db   514 HEKDFSLTEAPPGYDSVHGVSQTASVTTD------------------------FED 545

Qy   513 SEYLIYQESQCRLRYLLE 530
            ||:::|:: :|::::
Db   546 DEFVYKTNQVKMKYIIK 563
```

Search completed: August 29, 2002, 08:02:15
Job time: 292 sec

RT   "The nuclear protein PH5P of the inter-alpha-inhibitor superfamily: a
RT   missing link between poly(ADP-ribose)polymerase and the
RT   inter-alpha-inhibitor family and a novel actor of DNA repair?";
RL   FEBS Lett. 446:6-8(1999).

CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBUNIT: BINDS TO THE MAJOR VAULT PROTEIN (MPV). THE VAULT, A
CC   LARGE RIBONUCLEOPROTEIN COMPLEX, CONTAINS THE 100-kDA MPV AND 2
CC   PROTEINS OF 193 AND 240-kDA.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, BUT IS ALSO FOUND IN THE
CC   NUCLEUS, ASSOCIATED WITH MITOTIC SPINDLES.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; THE HIGHEST LEVELS ARE IN
CC   THE KIDNEY; ALSO DETECTED IN HEART, PLACENTA, LUNG, LIVER,
CC   SKELETAL MUSCLE, SPLEEN, LEUKOCYTES AND PANCREAS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF158255; AAD47250.1; -.
DR   EMBL; AF057160; AAC62491.1; -.
DR   EMBL; AL359763; CAC21562.1; -.
DR   EMBL; D79999; BAA11494.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR002035; vWFA.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00292; BRCT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation; Ribonucleoprotein.
FT   DOMAIN   1   94   BRCT.
FT   DOMAIN   256   569   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   876   1046   VWFA.
FT   DOMAIN   1562   1724   INTERACTION WITH THE MAJOR VAULT PROTEIN.
FT   DOMAIN   19   25   NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT   (POTENTIAL).
FT   DOMAIN   1237   1249   NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT   (POTENTIAL).
FT   CONFLICT   519   519   S -> P (IN REF. 1).
FT   CONFLICT   873   873   S -> N (IN REF. 2 AND 4).
FT   CONFLICT   897   897   Q -> E (IN REF. 2 AND 4).
FT   CONFLICT   899   899   A -> T (IN REF. 1).
FT   CONFLICT   936   936   A -> T (IN REF. 1).
FT   CONFLICT   1012   1012   A -> M (IN REF. 3).
FT   CONFLICT   1065   1065   V -> I (IN REF. 4).
FT   CONFLICT   1080   1080   A -> V (IN REF. 3 AND 4).
FT   CONFLICT   1108   1108   R -> L (IN REF. 3).
FT   CONFLICT   1265   1265   C -> R (IN REF. 2 AND 4).
FT   CONFLICT   1280   1280   G -> A (IN REF. 2 AND 4).
FT   CONFLICT   1328   1328   T -> P (IN REF. 3).
FT   CONFLICT   1331   1331   A -> A (IN REF. 3).
FT   CONFLICT   1394   1394   A -> S (IN REF. 3).
FT   CONFLICT   1459   1459   Y -> S (IN REF. 3).
FT   CONFLICT   1550   1550   P -> L (IN REF. 1).
FT   CONFLICT   1555   1555   V -> L (IN REF. 1).
FT   CONFLICT   1564   1564   T -> I (IN REF. 3).
FT   CONFLICT   1656   1656   P -> A (IN REF. 3).
SQ   SEQUENCE   1724 AA; 192587 MW; A009F3493460EDC CRC64;

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

```
                Copyright (c) 1993 - 2000 Compugen Ltd.
                         GenCore version 4.5
```

Run on: August 29, 2002, 08:01:28 ; Search time 117.38 Seconds
                                          (without alignments)
                                          785.537 Million cell updates/sec

```
Title:               US-09-701-586B-4
Perfect score:       2823
Sequence:            1 MAPKPKPWVQTEGPEKKGR.........EYLIYQESQCRLRYLLEVHL 533

Scoring table:       BLOSUM62
                     Gapop 10.0 , Gapext 0.5

Searched:            562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters:          562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    SPTREMBL_19:*
              1: sp_archea:*
              2: sp_bacteria:*
              3: sp_fungi:*
              4: sp_human:*
              5: sp_invertebrate:*
              6: sp_mammal:*
              7: sp_mhc:*
              8: sp_organelle:*
              9: sp_phage:*
             10: sp_plant:*
             11: sp_rodent:*
             12: sp_virus:*
             13: sp_vertebrate:*
             14: sp_unclassified:*
             15: sp_rvirus:*
             16: sp_bacteriap:*
             17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2823	100.0	533	4	Q96CG2	Q96cg2 homo sapien
2	2249.5	79.7	528	11	Q91YR6	Q91yr6 mus musculu
3	826	29.3	612	5	Q9TX06	Q9tx06 dictyosteli
4	694	24.6	635	10	081294	081294 arabidopsis
5	670.5	23.8	653	10	050017	050017 zea mays (m
6	669.5	23.7	607	12	Q9PS82	Q9ps82 gallus gall
7	653	23.1	1014	11	Q921K2	Q921k2 mus musculu
8	653	23.1	607	13	Q9PS81	Q9ps81 xenopus. na
9	636.5	22.5	945	5	Q9N4H4	Q9n4h4 caenorhabdi
10	629	22.3	593	5	Q9TX05	Q9tx05 drosophila
11	598.5	21.2	983	10	Q9ZP54	Q9zp54 arabidopsis
12	598.5	21.2	1009	10	Q9SJW4	Q9sjw4 arabidopsis
13	578	20.5	969	10	024570	024570 zea mays (m
14	578	20.5	980	10	Q9ZSV1	Q9zsv1 zea mays (m
15	534	18.9	727	5	Q9XUA5	Q9xua5 caenorhabdi
16	306.5	10.9	2276	5	Q9TXQ1	Q9txq1 caenorhabdi
17	302	10.7	815	10	Q9SWB4	Q9swb4 glycine max
18	284	10.1	815	10	Q9FK91	Q9fk91 arabidopsis
19	138.5	4.9	1327	4	O95271	095271 homo sapien
20	137	4.9	1166	4	Q9H2K2	Q9h2k2 homo sapien
21	137	4.9	1265	4	Q9HAS4	Q9has4 homo sapien
22	134	4.7	1156	16	O66878	o66878 aquifex aeo
23	127	4.5	363	4	Q9H8R9	Q9h8r9 homo sapien
24	125	4.4	1181	5	Q9X237	Q9x237 drosophila
25	125	4.4	1181	5	Q9VBP3	Q9vbp3 drosophila
26	124.5	4.4	429	11	Q924M2	Q924m2 mus musculu
27	122.5	4.3	848	4	Q9Y2K0	Q9y2k0 homo sapien
28	120.5	4.3	419	4	Q9H9X9	Q9h9x9 homo sapien
29	120	4.3	1219	5	Q9NJ23	Q9nj23 aequipecten
30	120	4.3	1229	5	Q9NJ22	Q9nj22 aequipecten
31	120	4.3	1243	5	Q9NJ21	Q9nj21 aequipecten
32	120	4.3	1253	5	Q9NJ20	Q9nj20 aequipecten
33	120	4.3	1951	5	Q17042	Q17042 aequipecten
34	119.5	4.2	927	11	Q91V71	Q91v71 mus musculu
35	119.5	4.2	946	11	Q91XT8	Q91xt8 mus musculu
36	119.5	4.2	1021	11	Q91XU9	Q91xu9 mus musculu
37	119.5	4.2	1052	11	Q91XT7	Q91xt7 mus musculu
38	119	4.2	2167	10	Q9SSU1	Q9ssu1 chara coral
39	119	4.2	2182	10	Q9LW97	Q9lw97 chara coral
40	117	4.1	261	4	Q9H8F2	Q9h8f2 homo sapien
41	116.5	4.1	1435	2	Q03291	Q03291 saccharomyc
42	116	4.1	429	2	Q9LAX7	Q9lax7 streptococc
43	115.5	4.1	523	11	Q9DA63	Q9da63 mus musculu
44	115.5	4.1	744	4	Q9BXY9	Q9bxy9 homo sapien
45	114.5	4.1	1214	10	Q9FHI5	Q9fhi5 arabidopsis

ALIGNMENTS

```
RESULT   1
Q96CG2
ID   Q96CG2      PRELIMINARY;      PRT;     533 AA.
AC   Q96CG2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 60.1 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014260; AAH14260.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   533 AA;   60089 MW;   6296A0E439CC7767 CRC64;

Query Match             100.0%;  Score 2823;  DB 4;  Length 533;
Best Local Similarity   100.0%;  Pred. No. 2.3e-208;
Matches  533;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 MAPKPKPWVQTEGPEKKGRQAGREEDFFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAPKPKPWVQTEGPEKKGRQAGREEDFFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60

Qy   61 QYYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 QYVEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 120

Qy  121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 180

Qy  181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240
```

```
Db  181  QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEAL  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241  KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA  300

Qy  301  VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK  360

Qy  361  VNQGEEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  VNQGEEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS  420

Qy  421  KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDT  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDT  480

Qy  481  QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL  533
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL  533


RESULT 2
Q91YR6
ID   Q91YR6      PRELIMINARY;    PRT;    528 AA.
AC   Q91YR6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 59.4 KDA PROTEIN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014870; AAH14870.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE  528 AA;  59413 MW;  82FB0C498EB5F74 CRC64;

Query Match        79.7%;
Best Local Similarity  80.7%;      Score 2249.5;  DB 11;  Pred. No. 2.4e-164;
Matches  431;  Conservative  37;  Mismatches  59;  Indels  .  7;  Gaps  4;

Qy   1  MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT  60
        |||| :|.|||| .||:| | | : | .| :|||||||:| :|| ||||||:||:| |||
Db   1  MAPKRKASVQTEG--SKKORQGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI  58

Qy  61  QVYEDYNCTLNQTNIENNNKFYIYQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED  120
        ||:||| ||||||||.|||||||||||::.|||||:|||||||||||||||:|||| |||
Db  59  QVHEDYDCTLNQTNIGNNNKFYIYQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED  117

Qy 121  AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV  180
        |||||||||.|||||:|.||||||.:|.|||||||||.|.||||||||||.|||||   |
Db 118  AKKDFEKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEAESQEAVVKVDSGPVRTV---V  174

Qy 181  QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEAL  240
        :||||||||.|||||||||||||||:| ||:||| ::|||| |||||||||||||||:|:
Db 175  KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKMNLGKLTKQQIARGFEALEALEEAM  234

Qy 241  KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA  300
        | || |||||||||| |||||||||| |: | ||||::||||||||||||||||| |||
Db 235  KNPTGDGGQSLEELSSCFYTVIPHNFGRSRPPFINSPDVLQAKKDMLLVLADIELVQTLQA  294

Qy 301  V-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIW  359
        . .: :||:||||||||||||||:||||||||| ||:::|||| | || :|||| ||:|:
Db 295  APGEEEEKVEEVPHPLDRDYQLLRCQLQLLDSGESESEYKAIQTYLKQTGNSYRCPNLRHVW  354
```

```
Qy  360  KVNQGEEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASEN  419
         ||||:|.|.||||||||||||:||||||.|||||||||||||||||||||||||||||||
Db  355  KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASEN  414

Qy  420  SKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP  479
         |||||||  | || |:||||||||||||:|||| |:|||||||||||||||||| |||||
Db  415  SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP  474

Qy  480  TQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL  533
         .||:ELELDGQ:VVVPQGQPV   :SSSFSQS  LIY:ESQCRLRYLLE:HL
Db  475  AQDIELELDGQPVVVPQGQPVPQCPSFKSSSFSQSFLIYKESQCRLRYLLEIHL  528


RESULT 3
Q9TX06
ID   Q9TX06      PRELIMINARY;    PRT;    612 AA.
AC   Q9TX06;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
SQ   SEQUENCE  612 AA;  69241 MW;  CB340F7A88FF2364 CRC64;

Query Match        29.3%;
Best Local Similarity  38.5%;      Score 826;  DB 5;  Pred. No. 5.8e-55;
Matches  217;  Conservative  79;  Mismatches  194;  Indels  74;  Gaps  19;

Qy   3  PKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRV------DPTCPL  54
        | :|| ||| |::| | | : | .|.:.| :|||
Db  86  PTKKTKTNTSEAETK---TAASDLDDSSSSESE-----DEKNQISVKIKGRAANDPHFPD  136

Qy  55  SSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ  109
        | : ::|. :    :.  ||||||  :|||.|||||:  ::.  .|.|||| :|  |:|
Db 137  SR---XHVYENGKDVYDATLNQTEISQNNNKXYIIQLLEADDGSSYSVWNRWGEGLKGQ  193

Qy 110  SKINHFTR--LEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVK  167
        | |:| .:  | ||| | :. || |||| . | : .. .|||.:::|  :.  .  :
Db 194  SSRKDFGKGGLNQAISLFCSKFYEKTKNTFTDRANFKKVAGKYDMIELDYSTDS-----K  248

Qy 168  VDRG----PVRTVTKRV----QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLG  219
        . .|     :  .| ::     .||||| .::::   :|.|. ..:.  ::  :||:|||
Db 249  PKNGASTTATTTTKKVVEHKKECSLDERVQELVKLIFDVKMMERTMTEAKYDLKMPLG  308

Qy 220  KLSKQQIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELL  279
        |||::||::|. .|.::|:.:.| :. ||||||||| ||| ||| ||.| .| ||   |
Db 309  KLSKNQITKGYLVLKQIEDVMGGKS--GESLSTLSSRFYTIIPHAFGMSVPPVINFNQML  366

Qy 280  QAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVI  339
        :.|::. :: | ||: ..  . .:  . :    ::.:    |: .|  :  :|: |||:|
Db 367  IEKMNMLQNLADIEIATNIIKDSESDES-----NLELHYAKLKTDIQPLDENSCEYKNI  421

Qy 340  QTYLEQTGSNHRCPTLQHIWKVNQGEEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGL  399
        ||| :.| .:.. |.:.:::|: .:|  ..::. :: .|||||||| : :. .: || ||
Db 422  LLYVKNTYQGGKKPTIVNIFKIDRDGEADRYKTKKHLGNRKLLWHGSRLINYASIISQGL  481
```

```
Qy    400  RIMPH----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY-----MFLGEVALGREH  450
                ||| |:  | ||:||:||| | ||| |  ||    ||:  ||:||:||:
Db    482  RIAPPEAPVSGYRFGKGVYFADCMSLSANY------CRTVGFLIDFCMLLGDVALGKTA  534

Qy    451  HINTDNPSLKSPPPGFDSVIARGHTEPDPT---QDTELELDGQQVVVPQGQPVPCPEFSS  507
                ||||  ||:::| |||||  ::  ::||        ||||:|||||| || ||  :|
Db    535  DLARDT-YMEKFQPNSHSTWALGTVEPDPKVFFQDTE------GCITPYGQMIPSQHKGV  587

Qy    508  STFSQSEYLIYQESQCRLRYLLEV  531
                | : | :| ||  || ||| |
Db    588  SCY-EHQYVVVDVAQVHLKYLLQL  610


RESULT    4
081294
ID    081294;         PRELIMINARY;      PRT;     635 AA.
AC    081294;
DT    01-NOV-1998 (TrEMBLrel. 08, Created)
DT    01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT    01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE    T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE).
GN    T14P8.19 OR AT4G02390.
OS    Arabidopsis thaliana (Mouse-ear cress).
OC    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC    eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX    NCBI_TaxID=3702;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=CV. COLUMBIA;
RA    WASHU;
RT    "The A. thaliana Genome Sequencing Project.";
RL    Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN    [2]
RP    SEQUENCE FROM N.A.
RC    STRAIN=CV. COLUMBIA;
RA    Kalicki J., Elliott G., Cloud J.;
RT    "The sequence of A. thaliana T14P8.";
RL    Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN    [3]
RP    SEQUENCE FROM N.A.
RC    STRAIN=CV. COLUMBIA;
RA    Waterston R.;
RL    Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN    [4]
RP    SEQUENCE FROM N.A.
RA    Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA    Mayer K.F.X.;
RL    Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN    [5]
RP    SEQUENCE FROM N.A.
RA    EU Arabidopsis sequencing project;
RL    Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AF069298; AAC19283.1; -.
DR    EMBL; AL161494; CAB80732.1; -.
DR    HSSP; P26446; 1A26.
DR    InterPro; IPR001290; PARP.
DR    InterPro; IPR004102; PARP_reg.
DR    InterPro; IPR003034; SAP.
DR    Pfam; PF00644; PARP; 1.
DR    Pfam; PF02877; PARP_reg; 1.
DR    Pfam; PF02037; SAP; 2.
DR    SMART; SM00513; SAP; 2.
KW    Transferase.
SQ    SEQUENCE   635 AA;  72017 MW;  E3F1CBE4D367A377 CRC64;

Query Match             24.6%;  Score 694;  DB 10;  Length 635;
Best Local Similarity   34.2%;  Pred. No. 8.3e-45;
Matches  185;  Conservative   97;  Mismatches 221;  Indels  38;  Gaps  18;

Qy     11  TEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRV-DPTCPLSSNPGTQVYED----  65
                 :  |: : |  :| :  ||:  ||  |  | ||      |:
```

```
Db    107  SNAPVKSSNDEAEDDNNGPEEEKEEKIVTATKGAAVLDQWIPDEIKSQYHVLQRGDDV  166

Qy     66  YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK  123
                |: ||||:: ::||||:|:||:| |:||  |||||||||||:||| |  ||| ||:||:

Db    167  YDAILNQTNVRDNNNKFFVLQVLESDSKKTYMVTRWGRVGVGQSKLDGPYDSWDRAIE  226

Qy    124  DFEKKFREKTKNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQP-  182
                | || :|||| |||:|| | ||   ||:  ||:|| || ||  :||

Db    227  IFTNKFNDKTKNWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNND---IPSSSSEVKPE  283

Qy    183  -CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQOIARGPEALEEEALK  241
                 :|| | | : :| :|  :|  ||:: | | || ||||:| :  :   | |:|

Db    284  QSKLDTRVKAKFISLICNVSMMAQHMMEIGYNANKLPLGKISKSTISKGYEVLKRISEVI  342

Qy    242  GPTDGGQSLEELSHFYTVIPHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ  299
                ||| || | |: | |||| |: | |||||   :|||| | |:||| || | ||:||||

Db    343  -DRYDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLL  401

Qy    300  AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ-  356
                : :|:    | |||:||| | : | |  :|| | |: |: |: |||| | | | | : :

Db    402  SVDPGLQ----DDPLYIHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVEIA  456

Qy    357  HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKG  412
                 : ::|||:|  |  ||:|||:|||| :| ||:|: |||:|| || |||     |||:|||

Db    457  QLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKG  516

Qy    413  IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR  472
                |:|  || |||| ||||:|||||| ||  ||||| :  |:: |: |   | |:  | |

Db    517  VFFADMFSKSANYCYA-NTGAND-GVLLLCEVALGDMNELLYSDYNADNLPPGKLSTKGV  574

Qy    473  GHTEPDPTQDTELELDGQOVVVPQGQPV--PCPEFSSSTFSQSEYLIYQESQCRLRYLLE  530
                | | || |:||| || |||||:||  || |::|| |||| ||  :||::|||||||||

Db    575  GKTAPNPSEAQTLE-DG--VVVPLGKPVERSC---SKGMLLYNEYIVYNVEQIKMRYVIQ  628

Qy    531  V  531

Db    629  V  629


RESULT    5
050017
ID    050017;         PRELIMINARY;      PRT;     653 AA.
AC    050017;
DT    01-JUN-1998 (TrEMBLrel. 06, Created)
DT    01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT    01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE    POLY(ADP-RIBOSE) POLYMERASE.
GN    PARP.
OS    Zea mays (Maize).
OC    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC    Panicoideae; Andropogoneae; Zea.
OX    NCBI_TaxID=4577;
RN    [1]
RP    SEQUENCE FROM N.A.
RA    Babyichuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA    O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT    "Higher plants possess two poly(ADP-ribose) polymerases.";
RL    Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AJ222588; CAA10888.1; -.
DR    HSSP; P26446; 1A26.
DR    InterPro; IPR001290; PARP.
DR    InterPro; IPR004102; PARP_reg.
DR    InterPro; IPR003034; SAP.
DR    Pfam; PF00644; PARP; 1.
DR    Pfam; PF02877; PARP_reg; 1.
DR    Pfam; PF02037; SAP; 2.
DR    SMART; SM00513; SAP; 2.
SQ    SEQUENCE   653 AA;  72995 MW;  5FD01923C4ABCD1D CRC64;

Query Match             23.8%;  Score 670.5;  DB 10;  Length 653;
Best Local Similarity   35.4%;  Pred. No. 5.5e-43;
```

```
Matches 170;  Conservative 93;  Mismatches 186;  Indels 31;  Gaps 14;

Qy    64 EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKINHFTRLED-A 121
          | : ||||||  : :|||||||| | :  ||||||| || |:
Db   188 EIYDATLNQTNVGDNNNKFYIIQVLESDAGGSPMVYNRWGRVGVRGQDKLHGPSPTRDQA 247

Qy   122 KKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ 181
          ||| || ||  :| :|: |  |||  |||  :  ::||   :::
Db   248 IYFEEGKFHNKTNNHWSDRKNFKCYAKKYTWLEMDYGETEKE---IEKG--SITDQIK   300

Qy   182 PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 241
          ||  | |||  ||||||||||| ||  ||| ||
Db   301 ETKLETRIAQFISLICNISMMKQRMVEIGYNAEKLPLGKLRKATILKGYHVLKRISDVIS 360

Qy   242 GPTDGGQSLEELSSHFYIVIPHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ 299
          : ||||||||| : |||||||| ||  |::  :  | ||||:|:|:||
Db   361 -KADRRHLEQLTGEFYIVIPHDFGFRKMREFIIDTPQKLKAKLEMVEALGEIEIATKL-  417

Qy   300 AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ-  356
             :   |   |:  :    | | :|: ||| :|||
Db   418 ---LEDDSSDQDDPLYARYKQLHCDFTPLEADSDEYSMIKSYLRNTHGKTHSGYTVDIV  473

Qy   357 HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKG 412
          |||: ::  | ||| ||||||||||||||| :||: ||||| ||  :    ||||||||
Db   474 QIFKVSRHGETERFQKFASTRNRMLLWHGSRLSNWAGIISQGLRIAPPEAPVTGYMFGKG 533

Qy   413 IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 472
          :||:: |||::  |  | ||  |||||:||||||| :||
Db   534 VYFADMFSKSANYCYASE--ACRSGVLLLCEVALGDMNELLNADYDANNLPKGKLRSKGV  591

Qy   473 GHTEPDPTQDTELELDGQQVVVPGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVH  532
          |:| ::  : :||| :||  |||  | |:||  : |
Db   592 GQTAPNMV-ESKVADDG--VVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMRYVLHVN 647


RESULT 6
Q9PS82
ID   Q9PS82      PRELIMINARY;      PRT;   607 AA.
AC   Q9PS82;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   607 AA;  68033 MW;  75F6EE1D30D8F402 CRC64;

Query Match             23.7%;  Score 669.5;  DB 13;  Length 607;
Best Local Similarity   33.2%;
Matches  186;  Conservative  100;  Mismatches  205;  Indels  69;  Gaps  24;

Qy     7 PW--VQTEGPE------------KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPT  51
```

```
Db    73 PWGAEVKTEHQEVAVDGKCSKPANMKSAGKVKEEQGPSKSEKKMKLTV--KGGAAVDPD 129
          || :|||  |  :       :   : :    |       : ||:
Qy    52 CPLSSNPGTQVYED----YNCTLNQTNIENNNKFYIIQLLQD-SNRFFTCWNRWGRVG  105
           |  |:| ::         :: :||||||||||||||  :::  ||:|||||||||
Db   130 SGLEDS--AHVFEEKGKIFSATLGLVDIVKGTNSYKLQLLEDDRESRYWV-FRSWGRVG 186

Qy   106 EV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQE 163
          |:  :|       |                  :    : :  :      ::  :
Db   187 TVIGSNKLEQMPSKEDAVEHFLNLYEEKTGNSWHSK-NFTKYPKKFYPLEIDYGQDEEAV 245

Qy   164 AVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSK 223
           : |: : :           :  : |  ::  | :     :||:|:|||||||||||||
Db   246 RKLTVSAG---TKSKLAKP------IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSK 296

Qy   224 QQIARGFEALEALEEALKGPTDGGQSLE-ELSSHFYTVIPHNFGHSQPPPINSPELLQA  281
           :|   :   | :: ::  :|||:   |  |:| ||| |||:||  :  || ||||: :
Db   297 RQIQSAYSILNEVQQAV--SDGGSESQIDLSNRFYTLIPHDFGMKKPPLLSNLEYIQA   353

Qy   282 KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQT 341
          |  ::|:|||||   :: :         :    :  :: :|:       : :|  :||:
Db   354 KVQMLDNLLDIEVAYSLLRGGNEDGDKD----PIDINYEKLRTDIKVVDKDSEEAKIIKQ 409

Qy   342 YLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSG 398
          | :   :  |   :::    ::  : || | |  :::|::|:|||||:   ::  | |:|
Db   410 YVKNTHAATHNAYDLKVVEIFRIEREGESQRXKPFKQLHNRQLLWHGSRTTNFAGIISQG 469

Qy   399 LRIMPH----SGGRVGKGIYFASENSKSAGYVIGMKC--GAHHVGYMFLGEVALGREHH  451
          ||  :     |||:|||||||: ::  ::        | | : | |:::|||||| :
Db   470 LRIAPPEAPVTGYMFGKGIYFRDMVSKSANY----CHTSQADPIGLILLGEVALGNMYE  524

Qy   452 INTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPGQPVPCPEFSSSTFS  511
          :     :|:  :|  :    | :| ||| |  ||| |   ||| | | |  |  :
Db   525 LKNASHITKL-PKGKHSVKGLGKTAPDPTATT--TLDG-VEVPLGNGI-STGINDTCLL  578

Qy   512 QSEYLIYQESQCRLRYLLEV 531
          |||::| :  ::|:|||:|
Db   579 YNEYIVYDVAQVNLKYLLKL 598


RESULT 7
Q9Z1K2
ID   Q9Z1K2      PRELIMINARY;      PRT;   1014 AA.
AC   Q9Z1K2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE)
DE   POLYMERASE).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC012041; AAH12041.1; -.
KW   Transferase.
SQ   SEQUENCE   1014 AA;  112721 MW;  5DCE68E4CB3F46EB CRC64;

Query Match             23.7%;  Score 669.5;  DB 11;  Length 1014;
Best Local Similarity   33.5%;  Pred. No. 1.3e-42;
Matches  186;  Conservative  99;  Mismatches  203;  Indels  67;  Gaps  22;

Qy     1 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRI------IRVDPTCP  53
          :|||   ||:|||      |:|:  ::           : ||:      :||||||
Db   494 VAPKGK-----SAAPSKKSKGAVKEEGVNKS---------EKRMKLTLKGGAAVDPDSG 538

Qy    54 LSSNPGTQVYED----YNCTLNQTNIENNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV-  107
           |  | ::           :: :||||||||||||  |:|| |  ||:|||||||||
Db   539 LEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYKLQLLEDDKESRYWIFRSWGRVGTVI  596
```

```
Qy    108  GQSKINHFTRLEDAKKDFEKKFREKTKNNWAAERDHFVSHPGKYTLIEVQ-AEDEAQEAVV  166
                |.|:      |||.| ||.|| |||.|||  :|:  |:  ||:   :||   |||
Db     89  PSSGPVAGKSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDSEL---EDSCHVLETGG---  142

Qy    597  GSNKLEQMPSKDAVEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGGDE--EAVK  653
                ||| :||    |||:  |  |:|| ||.| :   ||||.||||| |. ||  |||
Db     63  YEDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEV-GQSKINHFTRLED  120

Qy    167  KVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQI  226
               |: |   |: :.|||:  |||  ||:||||||:|||||||||||||||||||||||:||
Db    143  -KIFSATLGLVDITRGTNSYYKLQLIEHDRSRYWVFRSWGRVGTVIGSKKLEEMSSKED  201

Qy    654  KLTVKP-GTKSKLPKP------VQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQI  706
              |: |    |:      |:.||    |.|:  ||||||||||||||||||:||||||:||
Db    202  AIEHFLNLYQDKTGNAW-HSPNFTKYPKKFYPLEIDYGQE-EDVVKKLSVG-AGTKSKLA  258

Qy    227  ARGFEALEALEEALKGPTDGGGSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDML  286
             ||:  :| ::| ||   ||||||:||| |:|||||:::||||||||||||||:||||||||:|||||
Db    181  QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQIARGFEALEALEEAL  240

Qy    707  QAAYSILSEVQQAVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEML  765
                    |:  :|:    |::|: |||||||:||:||:||||||||||:|||||:||||||||||
Db    259  KP------VQELIKLIIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIOSAYSILSOVQQAV  312

Qy    287  LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT  346
                |   |||:          :||| ||||||||||||:  ||| :| |||||||:|||:|||||||:||
Db    313  ------SESLSEARLLDLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVA  366

Qy    347  -GSNHRCPTLQ--HIWKVNQEGEEDRRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP  403
              |||||||:||   |||||||||:||||||:||||:||||:|||||||||||||||||||||||||:
Db    367  YSLLRGGADDGEK------DPIDVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNA  420

Qy    404  H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDN  456
                    |||||| ||||||  |||||||||||:||||      :| |||||||||:|||||||
Db    421  YDLEVLEIFKIDREGEYQRYKPFQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTG  480

Qy    882  PEAPVTGYMFGKGIYFADMVSKSANY----CHTSQGDPIGLILLGEVALGNMYELKHAS  936
              |:         |||||||||  ||:  |||||    || ||  ||||:| :
Db    407  GRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGF  466

Qy    457  PSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYL  516
             || |   | | | | | ||:| |     ||||||:
Db    481  YMFGKGIYFADMVSKSANYCHAMP--GSPIGLILLGEVALGNMHELKAASQITKL-PKGK  537

Qy    937  -HISKLPKGKHSVKGLGKTTPDPSASITLE----GVEVPLGTGIP-SGVNDPCLLYNEYI  990
               | | || ||:  |||:.| |||:      |: ||||: | |
Db    467  DSVIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLR  526

Qy    517  IYQESQCRLRYLLEV  531
                 :| :|:||:|::|
Db    527  YLLEV  531

Qy    991  VYDIAQVNLKYLLKL 1005
                :| ||||||:|||
Db    593  YLLKL  597
```

RESULT 8

```
ID   Q9PS81       PRELIMINARY;       PRT;   607 AA.
AC   Q9PS81;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Xenopus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae.
OX   NCBI_TaxID=8353;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE  607 AA;  67496 MW;  54CDEBBE22079886 CRC64;
```

Query Match   23.1%;   Score 653; DB 13; Length 607;
Best Local Similarity 32.5%; Pred. No. 1.1e-41;
Matches 177;   Conservative 102; Mismatches 214; Indels 52; Gaps 20;

```
Qy     3  PKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQV  62
            |  ::  || :   |:     |  |:  | | ||  |  |  |
```

RESULT 9

```
ID   Q9N4H4       PRELIMINARY;       PRT;   945 AA.
AC   Q9N4H4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 108.0 KDA PROTEIN.
GN   Y71F9AL.18.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Bradshaw-Cordum H., Scott K., Graves T.;
RT   "The sequence of C. elegans cosmid Y71F9AL.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024200; AAF36011.1; -.
```

DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; PARP_2n_FINGER_2; 1.
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   945 AA; 108006 MW; 1D0A62C954BC6AD9 CRC64;

Query Match   22.5%; Score 636.5; DB 5; Length 945;
Best Local Similarity   34.0%; Pred. No. 3.9e-40;
Matches 167; Conservative 84; Mismatches 189; Indels 51; Gaps 15;

Qy   66 YNCTLNQTNIENNNKFYIIQLLQDSNR-FFTCWNRWGRVG-EVGQSKINHFTRLEDAKK 123
              | ||||| | |::| |||||: | :: :: |||||| :: || || :|
Db   477 YQATLSFTDLTQNKNSYYKIQLLKDDQRENYYVFRSWGRVGTEVGGNKHESYSNSNEAIL 536

Qy   124 DFEKKFREKTKNKNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPC 183
              ||: |:: |: :||:| || | :: || : | ||: ||| : || || ::|||
Db   537 KFQDVFHEKTKNDWIYRKHFKRKMPGMFSYVETDYSEFAQITDTEITPG----SKTLLPK 591

Qy   184 SLDPATQKLITNIFSKEMPKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEAL-KG 242
              :| |||||:|:: |||: | :||| |||||||||||||||| ::||
Db   592 SV----KEVMSIFDVENMKSALKSFEMDVNKMPLGRLSHNQINLAFEVLNDISDLLVKL 647

Qy   243 PTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVS 302
              | ||| :|| ||| ||| |||||| :: :| ::| || | :|| |:|
Db   648 PIDASRIL-DFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNMLNALLDIKFAYDQISGG 706

Qy   303 EQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPT-LQHIWK 360
              | ::|| || : | |::|:|| || : :| :|| |: || |||| ||: | ||
Db   707 DVPASTSLGIDPVDINYQKLKCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVELIDILK 766

Qy   361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGIYFA 416
              :::: | : :||||||||||||||: |: | : : :| || ||||||||
Db   767 LNRDNESSKFKRH--IGNRRLLWHGSGKNFAGILGQGLRIAPPEAPVSGYMFGKGVYFA 824

Qy   417 SENSKSAGYVIGMKCGAHHVGYMFLGEVALGR-EHHINTDNPSLKSPPPGFDSVIARG-- 473
              || | | || :: :: ||:: : || : :| :: || |||:|
Db   825 DMFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSVQGLGRQ 881

Qy   474 ------HTEPD----PTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQ 522
                   ||||| || | |:| || || |||||| |
Db   882 CPREIGSYNHPDGYTIPLGLTYMQLQGKQDV---------------DYHLLYNEFIVYDVDQ 928

Qy   523 CRLRYLLEVHL 533
              :::|:|::
Db   929 IQLKLVRVKM 939

RESULT 10
Q9TX05
ID   Q9TX05   PRELIMINARY;   PRT;   593 AA.
AC   Q9TX05;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Drosophila sp. (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7242;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9607847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "on the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";

RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   593 AA; 68018 MW; D9BA37E38B8E7CCD CRC64;

Query Match   22.3%; Score 629; DB 5; Length 593;
Best Local Similarity   32.1%; Pred. NO.7.4e-40;
Matches 167; Conservative 100; Mismatches 196; Indels 58; Gaps 19;

Qy   37 KAIPAEKRI------IRVDPTCPLSSNPGTQVYED----YNCTLNQTNIENNNNKFYIIQL 87
              | ||:| ||| :|::   :|||||||||:|||||||||
Db   99 KSMPVSRTFKVKDGLAVDPDSGLEDI--AHYVDSNNKYSVVLGLTDIQRNKNSYYKVQL 156

Qy   88 LQ-DSNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFREKTKNWAERDHFVS 145
              || || | ||||||||| ||||||||||:|:|| || ||| |:: || : |: :
Db   157 LKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVK 216

Qy   146 HPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNT 205
              || : |: : :|| |: :|| ||||: || ::
Db   217 RTGRMYPIEIQYDDD--QKLVKHE---------SHFFTSKLEISVQNLIKLIFDIDSMNKT 266

Qy   206 MALMDLDVKKMPLGKLSKQQIARGFEALEALEEEALKGPTDGGQSLEELSSHFYTVIPHNF 265
              : |:|||||||||||||| || ||| :| | ::| :|| | |||||||||||
Db   267 LMEFHIDMDKMPLGKLSAHQIOSAYRVKEIYNVLECGSNTAK-LIDATNRFYTLIPHNF 325

Qy   266 GHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQ 325
              | : :   : |:: :|: | ||| | || : : || :| || | | ||||| ||
Db   326 GVQLPTLIETHQQIEDLRQMLDSLAEIEVAYSI---IKSEDVSDACNPLDNHYAQIKTQ 381

Qy   326 LQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLL 382
              | | || | ||: || | |||| | | ||||| : :|::| | || : :
Db   382 LVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARREKPKKLHNRKLL 441

Qy   383 WHGTNMAVVAAILTSGLRI----MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY 438
              |||: |: : : : :| | : || ||||| | | || || :: :: |
Db   442 WHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFADMVSKSANYCCTSQ--QNSTGL 499

Qy   439 MFLGEVALG------REHHIN--TDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQ 490
              || | :| :: | ||:||||| | : |
Db   500 MLLSEVALGDMMECTSAKYINKLSNNK------------HSCFGRGRTMPDPTK-SYIRSDG- 548

Qy   491 QVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEV 531
              : :||:|| || ||| | | | :   :|| ::|
Db   549 -VEIPYGETI-TDEHLKSSLLYNEYIVYDVAQVNIQYLFRM 587

RESULT 11
Q9ZP54
ID   Q9ZP54   PRELIMINARY;   PRT;   983 AA.
AC   Q9ZP54;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).
GN   PARP-1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LANDSBERG ERECTA;
RA   Doucet-chabeaud G., Kazmaier M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ131705; CAA10482.1; -.

DR   HSSP; P26446; 1A26.
KW   Transferase; Glycosyltransferase; NAD.
FT   CHAIN   2   983   POLY(ADP-RIBOSE) POLYMERASE.
SQ   SEQUENCE   983 AA; 111232 MW; 468E12A8EF1B6F4F CRC64;

Query Match
Best Local Similarity   31.1%;   Score 598.5; DB 10; Pred. No. 3.4e-37;
Matches 171; Conservative 101; Mismatches 199; Indels 79; Gaps 23;

Qy   17 KKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQ----VYED----YNC 68
             || | :: | |:| :: . :   ::|:|:: |:| ||
Db   472 KKQRKLPFDKYKIEDTSESLVTVKVGR------SAVHEASGLQEHCHILEDGNSIYNT 524

Qy   69 TLNQTNIENNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD 124
             ||: | | :: |:|||:::| ::   :|||| |||| :|: : ||
Db   525 TLSMSDLSTGINSYYILQIIQE-DKGSDCYVFRKWGRVGNKVEEMSK-SDAVHE 582

Qy   125 FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ 181
             |:::||:||:| : | :: :: :|:   :::
Db   583 FKRLFLEKTGNTWESWEQKTNFQKQPGKFLPLD----------------IDYGVNKQVAKK-E 628

Qy   182 P----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 237
             | :|::||:| | |::| :||:|:||| |||::: | :|||| | |||
Db   629 PFQTSSNLAPSLIELMKLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ 688

Qy   238 EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIE 293
             ||| ||||| ::| ||:||:|:| |:: | : : :: |:| |||::|||
Db   689 RLLTESDPQPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIE 744

Qy   294 LAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP 353
             :|:||:| ::|:|: | |: : ||: : | | ::|| | | |: ||: ||
Db   745 IASRI--VGFDVDSTES----LDDKYKKLHCDISPLPHDSEDYRLIEKYLNTT----HAP 794

Qy   354 T-------LQHIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRI---- 401
             | |:|:|:|::| ::||::| |||| |:| || :|:| | |:|:|
Db   795 THTEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPE 854

Qy   402 MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS 461
             | ||||||||||||::::||||| |:| || :| |:|: | |||: :: |: |:|
Db   855 APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGEIHEL-TKAKYMDK 911

Qy   462 PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES 521
             |||| : :| |::| :|::||| ||::||||| | : ||:   : :|| ||:
Db   912 PPRGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEYIVYDTA 967

Qy   522 QCRLRYLLEV 531
             | ||:||||
Db   968 QVKLQFLLKV 977

RESULT   12
Q9SJW4
ID   Q9SJW4;   PRELIMINARY;   PRT;   1009 AA.
AC   Q9SJW4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE POLY (ADP-RIBOSE) POLYMERASE.
GN   AT2G31320.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,

RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006593; AAD20677.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE 1009 AA; 114133 MW; CDE6E41CC2A3A2DB CRC64;

Query Match
Best Local Similarity   31.1%;   Score 598.5; DB 10; Pred. No. 3.5e-37;
Matches 171; Conservative 101; Mismatches 199; Indels 79; Gaps 23;

Qy   17 KKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQ----VYED----YNC 68
             || | :: | |:| :: . :   ::|:|:: |:| ||
Db   498 KKQRKLPFDKYKIEDTSESLVTVKVGR------SAVHEASGLQEHCHILEDGNSIYNT 550

Qy   69 TLNQTNIENNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD 124
             ||: | | :: |:|||:::| ::   :|||| |||| :|: : ||
Db   551 TLSMSDLSTGINSYYILQIIQE-DKGSDCYVFRKWGRVGNKVEEMSK-SDAVHE 608

Qy   125 FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ 181
             |:::||:||:| : | :: :: :|:   :::
Db   609 FKRLFLEKTGNTWESWEQKTNFQKQPGKFLPLD----------------IDYGVNKQVAKK-E 654

Qy   182 P----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 237
             | :|::||:| | |::| :||:|:||| |||::: | :|||| | |||
Db   655 PFQTSSNLAPSLIELMKLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ 714

Qy   238 EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIE 293
             ||| ||||| ::| ||:||:|:| |:: | : : :: |:| |||::|||
Db   715 RLLTESDPQPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIE 770

Qy   294 LAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP 353
             :|:||:| ::|:|: | |: : ||: : | | ::|| | | |: ||: ||
Db   771 IASRI--VGFDVDSTES----LDDKYKKLHCDISPLPHDSEDYRLIEKYLNTT----HAP 820

Qy   354 T-------LQHIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRI---- 401
             | |:|:|:|::| ::||::| |||| |:| || :|:| | |:|:|
Db   821 THTEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPE 880

Qy   402 MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS 461
             | ||||||||||||::::||||| |:| || :| |:|: | |||: :: |: |:|
Db   881 APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGEIHEL-TKAKYMDK 937

Qy   462 PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES 521
             |||| : :| |::| :|::||| ||::||||| | : ||:   : :|| ||:
Db   938 PPRGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEYIVYDTA 993

Qy   522 QCRLRYLLEV 531
             | ||:||||
Db   994 QVKLQFLLKV 1003

RESULT 13
O24570
ID   O24570   PRELIMINARY;   PRT;   969 AA.
AC   O24570;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babiychuk E., Cottrill·P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants posses two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ222589; CAA10889.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE 969 AA; 109128 MW; EB23AC62EEC14009 CRC64;

Query Match   20.5%; Score 578; DB 10; Length 969;
Best Local Similarity 30.6%; Pred. No. 1.2e-35;
Matches 166; Conservative 96; Mismatches 195; Indels 86; Gaps 21;

Qy   16 KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNI 75
            | ||| || ||| | : || :| | |||| | :: :|:| :::
Db   480 KVKGRSAVHESSGLQDTAHILE---------------------DGKSI---YNATLNMSDL 516

Qy   76 ENNNNKFYIIQLL-QDSNRFFTCWNRWGRVG--EVGQSKINHFTRLEDAKKDFEKKFREK 132
            | :||:|::: || : ||||| ||| ||| | ||| || || |:|| |:| ||
Db   517 ALGVNSYYVLQIIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEK 575

Qy   133 TKNNWAE--RDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPAT------LQ 189
            ||:||| | |:| |||:| :| || :|| || ||| |:| || :   :|
Db   576 TGNSWEAWECKTNFRKQPGRFYPLDVD------YGVKKAPKRKDISEMK-SSLAPQL 625

Qy   190 QKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEEAALKGPTDGGQS 249
            |||||| : |:| ||| |: : || || || :|| | | :||||| || | ||::
Db   626 LELMKMLFNVETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALA 685

Qy   250 LEE-----LSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQAV-SE 303
            ::   : :|:| | || : :: | | | | ::| | | || |:||:| :|
Db   686 VRESLIVAASNRFFTLIP----SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSD 741

Qy   304 OEKTVEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT------LQ 356
            ::: || | | | || | | : : || : | :| :|
Db   742 SDES------LDDKYMKLHCDITPLAHDSEDYKLIEQYL----LNTHAP THKDWSLELE 790

Qy   357 HIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAIITSGLRIMPH----SGGRVGK 411
            :: |::|| ::: | :: : | :: ||| :: | : : ||:: |: |
Db   791 EVFSLDRDGELNKYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGK 850

Qy   412 GIYFASENSKSAGYVIGMKCGA---HHVGIYMFLGEVALGREHHINTDNPSLKSPPPGFDS 468
            ||| :: || | | :| ||| || || | : | :::|
Db   851 GLYFADIVSKSAQY-----CVYDRNNPVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS 904

Qy   469 VIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSTFSQSEYLIYQESQCRLRYL 528
            :|:: || : |||::|| | : |||| || | || :|| ||:||:| :|:   :|
Db   905 TKGLGKTVP---LESEFVKRDDVVVPCGKPVP-SSIRSSELMNEYIVYNTSQVKMQFL 960

Qy   529 LEV 531
            :|:
Db   961 LKV 963

RESULT 14
Q92SV1
ID   Q92SV1   PRELIMINARY;   PRT;   980 AA.
AC   Q92SV1;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP)-RIBOSE POLYMERASE (EC 2.4.2.30).
GN   PARP1.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99026291; PubMed=9808734;
RA   Mahajan P.B., Zuo Z.;
RT   "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL   Plant Physiol. 118:895-905(1998).
EMBL; AF093627; AAC79704.1; -.
HSSP; P26446; 1A26.
InterPro; IPR001357; BRCT.
InterPro; IPR001290; PARP.
InterPro; IPR004102; PARP_reg.
InterPro; IPR003034; SAP.
InterPro; IPR001510; znf-PARP.
Pfam; PF00533; BRCT; 1.
Pfam; PF00644; PARP; 1.
Pfam; PF02877; PARP_reg; 1.
Pfam; PF00645; zf-PARP; 2.
ProDom; PD004675; Znf-PARP; 2.
SMART; SM00292; BRCT; 1.
PROSITE; PS50172; BRCT; 1.
PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase.
SQ   SEQUENCE 980 AA; 110475 MW; 9D8AED26BC37E5C1 CRC64;

Query Match   20.5%; Score 578; DB 10; Length 980;
Best Local Similarity 30.6%; Pred. No. 1.3e-35;
Matches 166; Conservative 96; Mismatches 195; Indels 86; Gaps 21;

Qy   16 KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNI 75
            | ||| || ||| | : || :| | |||| | :: :|:| :::
Db   491 KVKGRSAVHESSGLQDTAHILE---------------------DGKSI---YNATLNMSDL 527

Qy   76 ENNNNKFYIIQLL-QDSNRFFTCWNRWGRVG--EVGQSKINHFTRLEDAKKDFEKKFREK 132
            | :||:|::: || : ||||| ||| ||| | ||| || || |:|| |:| ||
Db   528 ALGVNSYYVLQIIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEK 586

Qy   133 TKNNWAE--RDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPAT 189
            ||:||| | |:| |||:| :| || :|| || ||| |:|
Db   587 TGNSWEAWECKTNFRKQPGRFYPLDVD------YGVKKAPKRKDISEMK-SSLAPQL 636

Qy   190 QKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEEAALKGPTDGGQS 249
            |||||| : |:| ||| |: : || || || :|| | | :||||| || | ||::
Db   637 LELMKMLFNVETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALA 696

Qy   250 LEE-----LSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQAV-SE 303
            ::   : :|:| | || : :: | | | | ::| | | || |:||:| :|
Db   697 VRESLIVAASNRFFTLIP----SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSD 752

```
Qy   304  QEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT------LQ  356
              :::    || |  |  :  :||:||: ||   || |||
Db   753  SDES------LDDKYMKLHCDITPLAHDSEDYKLIEQYL----LNTHAPTHKDWSLELE  801

Qy   357  HIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGK  411
              :||::|| :||:|| |||||||||||||||:||:|||||||| |||          ||  ||
Db   802  EVFSLDRDGELNKYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGK  861

Qy   412  GIYFASENSKSAGYVIGMKCGA---HHVGTMFLGEVALGREHHINTDNPSLKSPPPGFDS  468
              |:|||       |||  |  ||:  ||  || | | :||| ||
Db   862  GLYFADLVSKSAQY-----CYVDRNNPVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS  915

Qy   469  VIARGHTEPPDTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYL  528
              |  |      || |||:|| | |::|:||  | | |:||:|| :: |||:   ||  |:|
Db   916  TKGLGKTVP---LESEFVKWRDDVVVPCGKPVP-SSIRSSELMYNEYIVINTSQVKMQFL  971

Qy   529  LEV  531
              |:|
Db   972  LKV  974

RESULT   15
Q9XUA5
ID   Q9XUA5          PRELIMINARY;      PRT;    727 AA.
AC   Q9XUA5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC8.1 PROTEIN.
GN   AC8.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McMurray A.A.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z83097; CAB05448.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
SQ   SEQUENCE   727 AA;  82884 MW;  530ABA8E991FFEFD CRC64;

Query Match             18.9%;  Score 534;  Pred. No. 1.9e-32;
Best Local Similarity   28.6%;       DB 5;  Length 727;
Matches  162;  Conservative 104;  Mismatches 211;  Indels 90;  Gaps  20;

Qy    30  RSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQ  89
              ::|:|: |:  :| |                    ::|:|  |: |||
Db   182  KATGEYVEALAAKGG--STEPATPASASP-TPPEAETPVLSAEGSPESSNKRPASAEIIE  238

Qy    90  -------DSNRFF--TCWNRNGRVGEVGQSKI--------------------NHFTRLED  120
                      :|:|||  :|  | | :||  :
Db   239  IDGEGNPDENDFAKKRRMKKEARLMEVQKKRMKKQSDLLWEYRQIFERMPYTDNISILRE  298

Qy   121  AKKDF-------EKKFREKIKNNNAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGP  172
              :|||:|       |  |                         :: |:|
Db   299  NEQDIPEGHDRTAQDVFHEKTKNDWIYRKHFRKMPGMFSYVETDYSE-----FVGTNNGH  353
```

```
Qy   173  VRTVT--KRVQPCS---LDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA  227
              : :|           |  | ||   ||| | ||   ||  ||  :||  || | |  |:
Db   354  KKKITPGSKITPGSKTLLPKSVKEVVMSIFDVENMKSALKSFEIDVNKMPLGRLSHNQIN  413

Qy   228  RGFEALEAEEAL-KGPTDGGQSLEELSHFYTVIPHNFGHSOQPPINSPELLQAKKDML  286
              |   | || || ||||| ||| ||||||| |:||||||||| ||||| :||  ::||
Db   414  LAFEVLNDISDLLVKLPIDASKIL-DFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNML  472

Qy   287  LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT  346
              || : |    :| ::  |: :  |||  :||:|     | :|:|| |  ||::| :|:
Db   473  NALLDIKFAYDQISGGDVPASTSLSIDPVDINYRKLKCIMEPLQQGCDDWNMIHQYLKNT  532

Qy   347  -GSNHRCPT-LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH  404
               :| :||    |:  |||:||||||  |||||||||||||||||  | :| |||||||
Db   533  HGATHDLKVELIDILKVNRDNESSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPP  590

Qy   405  ----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGR-EHHINTDNPSL  459
                  || || ||||| |||||| ||||| |  |:|||:|| | | |:  :||||| |||
Db   591  EAPVSGYMFGKGVYFADMFSKSFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSR  647

Qy   460  KSPPPGFDSVIARG--------HTEPD---PTQDTELELDGQQVVVPQGQPVPCPEFS  506
              :: |||| |  |           |: |       |:|:|:| |  :|:|  | | |:
Db   648  QTLPAGFQSVQGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQGKQNV-----------  694

Qy   507  SSTFSQSEYLIYQESQCRLRYLLEVHL  533
              ::||||| ||||||||:||||||: |
Db   695  DYHLLYNEFIVYDVDQIQLKYLVRKM  721
```

Search completed: August 29, 2002, 08:01:31
Job time: 363 sec

us-09-701-586b-4.rspt

THIS PAGE BLANK (USPTO)

```
                GenCore version 4.5
            Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     August 29, 2002, 07:57:14 ; Search time 124.84 Seconds
                                          (without alignments)
                                          480.454 Million cell updates/sec

Title:      US-09-701-586B-6
Perfect score:  2854
Sequence:   1 MSLLFLAMAPKPKPWVQTEG.........EYLIYQESQCRLRYLLEVHL 540

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:   747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters:     747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  A_Geneseq_032802:*
        1:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
        2:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
        3:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
        4:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
        5:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
        6:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
        7:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
        8:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
        9:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
       10:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
       11:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
       12:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
       13:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
       14:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
       15:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
       16:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
       17:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
       18:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
       19:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
       20:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
       21:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
       22:  /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	2854	100.0	540	21	AAY51176	Human uterus type
2	2823	98.9	533	21	AAY51175	Human brain PARP3
3	2811	98.5	533	22	AAU29021	Human PARP-3 prote
4	2253.5	79.0	528	21	AAY51178	Murine PARP1 (shor
5	2241	78.5	533	21	AAY51177	Murine PARP1 (long
6	738	25.9	522	22	AAU29022	Mouse PARP-2 prote
7	729.5	25.6	583	22	AAB47029	hPARP2. Homo sapi
8	726.5	25.5	534	21	AAB42909	Human ORFX ORF2673
9	726.5	25.5	534	22	AAU29023	Human PARP-2 prote
10	725.5	25.4	534	22	AAU29020	Human PARP-2 prote
11	725	25.4	570	21	AAY51174	Human brain PARP2
12	725	25.4	570	22	AAB11480	Human brain poly-A
13	722	25.3	521	22	AAB60693	Human poly(ADP-rib
14	688.5	24.1	637	21	AAY68835	The poly(ADP-ribos
15	682.5	23.9	682	21	AAB93513	Human protein sequ
16	672	23.5	531	21	AAY58043	Human poly (ADP-ri
17	672	23.5	1014	22	AAU29019	Human PARP-1 prote
18	672	23.5	1014	22	AAB66296	Human tankyrase2 r
19	670.5	23.5	653	21	AAY68834	Human tankyrase2 r
20	670.5	23.5	1014	21	AAY49939	A poly(ADP-ribose)
21	670	23.5	1013	17	AAR99642	Human nuclear NAD+
22	664	23.3	1014	20	AAY33699	Poly(ADP-ribose) p
23	640.5	22.4	1063	22	AAB47032	Human poly(ADP-rib
24	638	22.4	557	22	ABB66431	Fusion protein PAR
25	578	20.3	969	21	AAY68833	Drosophila melanog
26	578	20.3	980	21	AAY68839	A poly(ADP-ribose)
27	475	16.6	982	20	AAY28464	Maize poly ADP-rib
28	472.5	16.6	379	22	AAU21687	Novel human neopla
29	447	15.7	360	21	AAB47030	N-terminal fragmen
30	410.5	14.4	1010	21	AAY68840	Fusion protein of
31	394	13.8	294	22	AAU20129	Human DNA repair a
32	394	13.8	294	22	AAU21810	Novel human neopla
33	390	13.7	287	22	AAB47031	C-terminal fragmen
34	288	10.1	1099	22	AAB66301	Human tankyrase2 e
35	267.5	9.4	227	22	AAU20130	Human DNA repair a
36	267.5	9.4	227	.22	AAU21811	Novel human neopla
37	258	9.0	1730	22	AAU33242	Novel human secret
38	256	9.0	1724	21	AAY54373	cDNA sequence enco
39	256	9.0	1724	22	AAB51022	Human minor vault
40	216.5	7.6	190	22	AAU21688	Novel human neopla
41	138.5	4.9	1327	21	AAB27212	Human tankyrase I
42	138.5	4.9	1327	21	AAY44402	Human tankyrase.
43	138.5	4.9	1327	22	AAB66279	Human tankyrase1 S
44	137	4.8	756	22	AAB66286	Human tankyrase2 c
45	137	4.8	784	22	AAB66285	Human tankyrase2 c

ALIGNMENTS

```
RESULT  1
AAY51176
ID   AAY51176 standard; Protein; 540 AA.
XX
AC   AAY51176;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human uterus type 2 PARP3 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage; PARP3.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;  99WO-EP03889.
XX
PR   05-JUN-1998;  98DE-1025213.
PR   01-MAR-1999;  99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AAZ44289.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
```

```
PT  conditions -
XX
PS  Claim 4; Page 62-64; 96pp; German.
XX
CC  This invention describes novel human and murine poly(ADP-ribose)
CC  polymerase (PARP) homologues, which are characterised by an amino acid
CC  sequence with a functional NAD^+-binding site and no zinc finger
CC  sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC  sequences, PARP homologues and antibodies are useful for analytic
CC  detection of PARP homologues and for identifying PARP effectors or
CC  binding partners, as well as for determining their effectiveness.
CC  PARP-binding partners are useful for the diagnosis or therapy of a
CC  disease condition, which is the result of a PARP protein, especially an
CC  energy deficiency, which may comprise tissue damage from cell death
CC  following necrosis or apoptosis. The disease condition may be chosen
CC  from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC  in particular neurotoxic disturbances, etc. This sequence represents the
CC  human PARP3 protein used in the method of the invention.
XX
SQ  Sequence 540 AA;

Query Match       100.0%;  Score 2854;  DB 21;  Length 540;
Best Local Similarity   100.0%;  Pred. No. 1.7e-252;
Matches  540;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  MSLLFLAMAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCP   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  mslllflamapkpkpwvqtegpekkgrqagreedpfrstaealkaipaekriirvdptcp  60

Qy   61  LSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKIN  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  lssnpgtqvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqgskin  120

Qy  121  HFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPV  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  hftrledakkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrgpv  180

Qy  181  RTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEAL  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  rtvtkrvqpcsldpatqklitnifskemfkntmalmdldvkkmplgklskqqiargfeal  240

Qy  241  EALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIE  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  ealeealkgptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladie  300

Qy  301  LAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  laqalqavseqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcp  360

Qy  361  TLQHIWKVNOEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGI  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  tlqhiwkvnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgi  420

Qy  421  YFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  yfasensksagyvigmkcgahhvgymflgevalgrehhintdnpslkspppgfdsviarg  480

Qy  481  HTEPDPTQDTELELDGQQVVVPQGGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  htepdptqdteleldgqqvvvpqggqpvpcpefssstfsqseyliyqesqcrlryllevhl  540

RESULT 2
AAY51175
ID   AAY51175 standard; Protein; 533 AA.
XX
AC   AAY51175;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Human brain PARP3 protein.
XX
```

```
KW  PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW  diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW  ischemic tissue damage; PARP3.
XX
OS  Homo sapiens.
XX
PN  WO9964572-A2.
XX
PD  16-DEC-1999.
XX
PF  04-JUN-1999;   99WO-EP03889.
XX
PR  05-JUN-1998;   98DE-1025213.
PR  01-MAR-1999;   99DE-1008837.
XX
PA  (BADI ) BASF AG.
XX
PI  Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR  WPI; 2000-087218/07.
DR  N-PSDB; AA244288.
XX
PT  Novel genes and proteins, antibodies and binding partners useful in
PT  diagnosis and therapy of energy deficiency associated disease
PT  conditions -
XX
PS  Claim 4; Page 57-59; 96pp; German.
XX
CC  This invention describes novel human and murine poly(ADP-ribose)
CC  polymerase (PARP) homologues, which are characterised by an amino acid
CC  sequence with a functional NAD^+-binding site and no zinc finger
CC  sequence motif, of general formula CX_2CX_2C (I). The nucleic acid
CC  sequences, PARP homologues and antibodies are useful for analytic
CC  detection of PARP homologues and for identifying PARP effectors or
CC  binding partners, as well as for determining their effectiveness.
CC  PARP-binding partners are useful for the diagnosis or therapy of a
CC  disease condition, which is the result of a PARP protein, especially an
CC  energy deficiency, which may comprise tissue damage from cell death
CC  following necrosis or apoptosis. The disease condition may be chosen
CC  from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC  in particular neurotoxic disturbances, etc. This sequence represents the
CC  human PARP3 protein used in the method of the invention.
XX
SQ  Sequence 533 AA;

Query Match        98.9%;  Score 2823;  DB 21;  Length 533;
Best Local Similarity    100.0%;  Pred. No. 1.1e-249;
Matches  533;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    8  MAPKPKPWVQTEGPEKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT   67
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  mapkpkpwvqtegpekkgrqagreedpfrstaealkaipaekriirvdptcplssnpgt   60

Qy   68  QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED  127
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  qvyedynctlnqtniennnkfyiiqllqdsnrfftcwnrwgrvgevgqgskinhftrled  120

Qy  128  AKKDFEKKFREKTKNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV  187
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrgpvrtvtkrv  180

Qy  188  QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL  247
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  qpcsldpatqklitnifskemfkntmalmdldvkkmplgklskqqiargfealealeeal  240

Qy  248  KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA  307
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  kgptdggqsleelsshfytviphnfghsqpppinspellqakkdmllvladielaqalqa  300

Qy  308  VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK  367
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  vseqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcptlqhiwk  360
```

```
Best Local Similarity  99.6%;  Pred. No. 1.4e-248;
Matches 531;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy      8 MAPKPKPWVQTEGPEKKGRQAGREEDFPRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 67
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mapkpkpwvqtegpekkgrqagreedfprstaealkaipaekriirvdptcplssnpgt 60

Qy     68 QYYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 qyyedynctlnqtniennnkkfyiiqllqdsnrfftcwnrwgrvgevggskinhftrled 120

Qy    128 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 187
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 akkdfekkfrektknnwaerdhfvshpgkytlievqaedeaqeavvkvdrapvrtvtkrv 180

Qy    188 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 247
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 qpcsldpatqklitnifskemfkntmalmdldvkkmplgklskqqiargfealealeeal 240

Qy    248 KGPTDGGQSLEELSSHFYTVIPHNFGHSQOPPPINSPELLQAKKDMLLVLADIELAQALQA 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 kgptdgggqsleelsshfytviphnfghsqppppinspellqakkdmllvladielaqalqa 300

Qy    308 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 367
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 vseqektveevphpldrdyqllkcqlqlldsgapeykviqtyleqtgsnhrcptlqhiwk 360

Qy    368 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 427
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 vnqegeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsggrvgkgiyfasens 420

Qy    428 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 487
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ksagyvigmkcgahhvgymflgevalgrehhintdnpslksppgfdsviarghtepdpt 480

Qy    488 QDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 qdteleldgqqvvvpqgqpvpcpefssstfsqseyliyqesqcrlryllevhl 533

RESULT  3
AAU29021
ID   AAU29021 standard; Protein; 533 AA.
XX
AC   AAU29021;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human PARP-3 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45590.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 105-107; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence  533 AA;

Query Match        98.5%;  Score 2811;  DB 22;  Length 533;
```

PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 71-73; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   murine PARP1 protein used in the method of the invention.
XX
SQ   Sequence 528 AA;

Query Match   79.0%;   Score 2253.5; DB 21; Length 528;
Best Local Similarity 80.9%; Pred. No. 1.8e-197;
Matches 432; Conservative 37; Mismatches 58; Indels 7; Gaps 4;

Qy   8 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 67
         |||| | |||| | || ||| ||||||::|| |:|||||:|| | |||
Db   1 mapkrkasvqteg--skkqrgqteeedsfrstaealraapadnrvirvdpscpfsrnpgi 58

Qy   68 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 127
         |:||||:||||||||||:|| ||||||||||||||||||||||||||||:||||||||||
Db   59 qvhedydctlnqtnignnnnkfyiiqlleegsrff-cwnrwgrvgevgqskmnhftcled 117

Qy   128 AKKDFEKKFREKIKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 187
         ||||:||| ||| :|| |||||||| | ||||||||| |: |:|| |||||| |
Db   118 akkdfkkkfwektknkweerdrfvaqpnkytlievggeaesqeavvkvdsgpvrtv---v 174

Qy   188 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 247
         :|||||||||||| ||||| | |:|||||||| |||||||||||:|||||||||||||||
Db   175 kpcsldpatqnlitnifskemfknamtlmnldvkkmpigkltkqqiargfealealeeam 234

Qy   248 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 307
         |||||| |||||||| ||| |||||||||||||||||| |:|||||||||||||| |
Db   235 knptgdggqsleelsscfytviphnfgrsrpppinspdvlqakkdmllvladielaqtlqa 294

Qy   308 V-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIW 366
         | :||||| ||||||||| ||||||||:||| | |:|||||||:||| ::| |:|:|
Db   295 apgeeekveevphpldrdyqllrcqlqlldsgeseykalqtylkqtgnsyrcpnlrhvw 354

Qy   367 KVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 426
         ||| |||| |||||||||||: ||||||||||||||||:||||||||||||||||||||||
Db   355 kvnregegdrfqahsklgnrrllwhgtnvavvaailtsglrimphsggrvgkglyfasen 414

Qy   427 SKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP 486
         |||||||| |||| |||||||||||| |||| | | ||||:|||| |||||| |:|
Db   415 sksagyvtmhcgghqvgymflgevalgkehhitiddpslkspppgfdsviargqtepdp 474

Qy   487 TQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 540
         || ||| | | ||||| | || | |::||| :|:||||: ||:|||||:|:|
Db   475 aqdieleldgqpvvvpqgqpvqcpsfksssfsqseyliykesqqrlrylleihl 528

RESULT 5
AAY51177
ID   AAY51177 standard; Protein; 533 AA.
XX
AC   AAY51177;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Murine PARP1 (long) homologue protein.

PT   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
PT   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
XX
KW   ischemic tissue damage; PARP1.
XX
OS   Mus sp.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999; 99WO-EP03889.
XX
PR   05-JUN-1998; 98DE-1025213.
PR   01-MAR-1999; 99DE-1008837.
XX
PA   (BADI) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI: 2000-087218/07.
DR   N-PSDB: AAZ44290.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 67-69; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   murine PARP1 protein used in the method of the invention.
XX
SQ   Sequence 533 AA;

Query Match   78.5%;   Score 2241; DB 21; Length 533;
Best Local Similarity 80.1%; Pred. No. 2.5e-196;
Matches 432; Conservative 37; Mismatches 58; Indels 12; Gaps 5;

Qy   8 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 67
         |||| | |||| | || ||| ||||||::|| |:|||||:|| | |||
Db   1 mapkrkasvqteg--skkqrgqteeedsfrstaealraapadnrvirvdpscpfsrnpgi 58

Qy   68 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 127
         |:||||:||||||||||:|| ||||||||||||||||||||||||||||:||||||||||
Db   59 qvhedydctlnqtnignnnnkfyiiqlleegsrff-cwnrwgrvgevgqskmnhftcled 117

Qy   128 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVK-----VDRGPVRT 182
         ||||:||| ||| :|| |||||||| | ||||||||| |: |:|| ||||||
Db   118 akkdfkkkfwektknkweerdrfvaqpnkytlievggeaesqeavvkvdsgpvrt 177

Qy   183 VTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA 242
         | |:|||||||||||| ||||| | |:|||||||| |||||||||||:|||||||||
Db   178 v---vkpcsldpatqnlitnifskemfknamtlmnldvkkmpigkltkqqiargfealea 234

Qy   243 LEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELA 302
         |||| |||||| |||||||| ||| |||||||||||||||||| |:||||||||||||
Db   235 leeamknptgdggqsleelsscfytviphnfgrsrpppinspdvlqakkdmllvladiela 294

Qy   303 QALQAV-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT 361
         | || |:: ||||||||||||| |||||||:||| | |:|||||||:||| ::| |:

```
Query Match                25.9%;  Score 738;  DB 22;  Length 522;
Best Local Similarity       34.3%;  Pred. No. 1.5e-58;
Matches  187;  Conservative 103;  Mismatches 191;  Indels  64;  Gaps  19;

Qy       23 KKKGRQA-GREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVY---EDYNCTL  77
             :||||||  ||| |      || | ||  :     |:| ||  ||
Db        5 qrkgpmaggkdadrtkdnrdsvktlllkgk-apvdpecaakigk-ahvycegddvydvml  62

Qy       78 NQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDAKKDFEKK 135
            ||||:|||  |:|||| || | |      :|| ||||| ||| |||||
Db       63 nqtnlgfnnnkyyliqlleddagrnfsvwmrwgrvgktgqhslvtcsgdlnkakelfgkk 122

Qy      136 FREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEA--QEAVVKVDRGPVRTVTKRVQ 188
            : :|||||| ||:|     |||    :|  :   ::
Db      123 fldktknnwedrenfekvpgkydmlqmdyaastqdesktkeeetlkpe--------- 170

Qy      189 PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 248
            |:||||||| | |  | ||  |: ::   ||::| :     |:  ::
Db      171 -sqldirvqelilklicnvqtmeemmiemkydtkrapigkltvaqikagyqslkkiedcir 229

Qy      249 GPTDGGGSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQAV 308
            | :|||| :||| :|  || | ||| |||| ||  ||| |    |:|
Db      230 a-gqhgralveacnefytriphdfglsippvirtekelsdkvkllealgdiei--alklv 286

Qy      309 SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHR--CPTLQHI 365
              :  :|   ||||| |||| || |: :|||  ||| |  :|||| :     ||  |
Db      287 kserqgle--hpldqhyrnlhcalrpldhesnefkvisqylgsthapthkdytmtlldv 343

Qy      366 WKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGKGIY 421
            | |||| ||||||||||||||||||||| |||| || ||||:|      ||||| |||
Db      344 fevekegekeafr--edlpnrmlwhgsrlsnvvglishglrvappeapitgymfgkgiy 401

Qy      422 FASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH 481
            |||||||||||||||||||||||||||||||||||||||||:||||:||   ||||:||
Db      402 fadmssksanycfasr--lkntglllsevalgqcnelleanpkagqllrgkhstkgngk 459

Qy      482 TEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSS------------TFSQSEYLIYQESQCRLR 533
             |||| :| || ||   ||   || | ||  ||                ||| |:|||| |||
Db      460 mapspa--------hfitlngstvplgpasdtglnpegytlnynefivyspnqvrmr 509

Qy      534 YLLEV 538
            ||:::
Db      510 yl1ki 514
```

RESULT 7

AAB47029

```
ID   AAB47029 standard; Protein; 583 AA.
XX
AC   AAB47029;
XX
DT   29-MAR-2001  (first entry)
XX
DE   hPARP2.
XX
CC   Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
CC   inflammation; ischaemic stroke; myocardial ischemia;
CC   infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
CC   gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
CC   endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
CC   toxic shock syndrome; multiple organ injury syndrome; vasculitis;
CC   hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
CC   eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
CC   chronic obstructive pulmonary disease; silicosis; reperfusion injury;
CC   pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
CC   bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
CC   scar tissue formation; atherosclerosis; systemic lupus erythematosus;
CC   autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
CC   graft versus host disease; allograft rejection; cystic fibrosis;
CC   chronic glomerulonephritis; inflammatory bowel disease;
CC   Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
CC   inflammatory dermatosis; contact dermatitis; atopic dermatitis;
CC   psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;
```

```
Db      295 qtlqaapgeeekveevphpldrdyqllrcqlqlldsgeseykaiqtylkgtgnsyrcpn 354
```

```
Qy      362 LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIY 421
            |:||||||:||  ||||||||||||||||||||||||:||||||||||||||||||||||
Db      355 lrhvwkvnregedrfqahsklgnrrllwhgtnvavvaailtsglrimphsggrvgkgiy 414

Qy      422 FASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH 481
            |||||||||||| :|||  | |   ||||||||||||:||||  |||| |||||||||||
Db      415 fasensksagyvttmhcgghqvgymflgevalgkehhitddpslkspppgfdsviargq 474

Qy      482 TEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSSTFSQSEYLIYQESQCRLRYLLEVHL 540
            |||||||||||| |||||||| | ||  ||||||::|||||||||||||| |||||| ||
Db      475 tepdpaqdieleldgqpvvvpqggppvqcpsfksssfsqseyliykesqcrlrylleihl 533
```

RESULT 6

AAU29022

```
ID   AAU29022 standard; Protein; 522 AA.
XX
AC   AAU29022;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Mouse PARP-2 protein.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Mus musculus.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45597.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13; Page 109-111; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer); cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia) , inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence  522 AA;
```

```
Qy   66  G-TQVY----EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SK  118
             : |||    ||||::|| :| |||||:::||| : | ||||||||:::||::||
Db  103  gkahvycegndvydvmlngtnlqfnnnkyyliqlleddagrnfsvwmrwgrvgkmqqhsl  162

Qy  119  INHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEAQEAVVK  174
             :  | ||||||| :|:|| :|:|| | ||  :: : ||||
Db  163  vacsgnlnkakelfqkkfldktknnwedrekfekvpgkydmlqndyatntqdeeetkkee  222

Qy  175  VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA  234
             :  |:: :  | | :::  :      ||:|  | ||  ||||:| ||
Db  223  slksplkpesq-----ldlrvqeliklicnvqameenmmemkyntkkaplgkltvaqik  276

Qy  235  RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL  294
             :  ||||||||||| |  |  :|: :  |||  : ||  |  | :|||::|| :
Db  277  agyqslkkiedcira-gqhqralmeacnefytriphdfglrtppllrtqkelsekiqile  335

Qy  295  VLADIELAQALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQYTYLEQT  353
             |||| | :|:|   |||| : ::   | |||  |||:|:|:||  ||| || |:|| | |
Db  336  algdieialkl-----vktelgspehpldqhyrnlhcalrpldhesyefkvisgylgst  389

Qy  354  GS--NHRCPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP  410
             :   | |||:  |||| : :| : :|| | |||  :| ||||| :  ||||
Db  390  hapthsdytmtlldlfevekdgekeafr--edlnrmllwhgsrmsnwvgilshglriap  447

Qy  411  H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL  466
             |     :  |  :  |  || :|  :|  :  :  | |:|| ||: | : || |::
Db  448  peapitgymfgkglyfadmssksanycfasr--lkntglllsevalgqcnelleanpka  505

Qy  467  ------------------------KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVPQGQPVPCPE  511
                                     || | | | | :  | | |  | : : :  |
Db  506  egliqqkhstkglgkmapssahfvtlngstvplgpasdt------glnpdgy------  552

Qy  512  FSSSTFSQSEYLIYQESQCRLRYLLEV  538
             : :| ||| :| :|:|||||
Db  553  ---tlnyneyivynpnqvrmryllkv  575
```

```
RESULT   8
AAB42909
ID   AAB42909 standard; Protein; 534 AA.
XX
AC   AAB42909;
XX
DT   08-FEB-2001   (first entry)
XX
DE   Human ORFX ORF2673 polypeptide sequence SEQ ID NO:5346.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; antipsoriatic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
```

```
Qy   22  EKKKGRQAG------REED-----PFRSTA-----EALKAIPAEKRIIRVDPTCPLSSNP  65
             |  || ||      ||||      :|:|| |    |:|| |  :| ||
Db   46  eskmpvaggkankdrtedkqdgmpgrswaskrvsesvkalllkgk-apvdpec--takv  102
```

PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000;   2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA., Leach M;
XX
DR   WPI; 2000-602362/57.
DR   N-PSDB; AAC77118.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 11; Page 4522-4524; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence   534 AA;

Query Match   25.5%;   Score 726.5;   DB 21;   Length 534;
Best Local Similarity 35.1%;   Pred. No. 1.8e-57;
Matches 196; Conservative 99; Mismatches 187; Indels 77; Gaps 21;

Qy   20 GPEKKKGRQAGREED-PFRSTA----EALKAIPAEKRIIRVDPTCPLSSNPG-TQVY-- 70
             |:| |||| ::: ||||| | ::: |:|||| :: :| ||
Db   5 ggkankdrtedkqdgmpgrswaskrvsesvkalllkgk-apvdpec--takvgkahvyce 61

Qy   71 --EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLE 126
            |||||| ::|||| |::|:||::| |:|: ||||||||||| ::|| |
Db   62 gndvydvmlnqtnlqfnnnkyyliqllleddaqrnfsvwmrwgrvgkmgqhslvacsgnln 121

Qy   127 DAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEAQEAVVKVDRGPVRT 182
           ||:||| |||||||||| ::|||||| || ::: || :||| |||||||||:
Db   122 kakelfqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkp 181

Qy   183 VTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA 242
            || ||||| || |:| | ||||||||||||||||||||| || |:||::::
Db   182 esq-----ldlrvgeliklicnvgameemmmemkyntkkaplgkltvaqikagyqslkk 235

Qy   243 LEEALKGPTDGGQSLEELSSHFYTVTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELA 302
            :|:|||: |||| |||: |||| ||||||| |:|| |: |||||||||||||||||:|
Db   236 iedcira-gqhgraimeacnefytriphdfgirtpplirtqkelsekiqillealgiela 294

Qy   303 QALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHR 358
            :||: :|| ||:|| |||| :|:|||:||||| | :|||| ::|||: ---NHR
Db   295 lkl-----vktelgspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdy 348

Qy   359 CPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGG 414
            :: ||| | :: |||: | |||:|||:|||:| | || :||| | ||
Db   349 tmtlldlfevekdgekeafr--edlhnrmllwhgsrmsnwvgilshglriappeapitgy 406

Qy   415 RVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL------- 466
             |||||| |||||||||| ||||| || || ||||| |||| : : ||
Db   407 mfgkgiyfadmssksanycfasr--lkntglllisevalgqcnelleanpkaegllqgkh 464

Qy   467 ------KSPPPGFDSVIARGHTEP-DPTQDTELELDGGQVVVPQGOPVPCPEFSSSTFSQ 519
                      :| ||||| :: || | || |::||| ||
Db   465 stkglgkmapssahfvtingstvplgpasdt--------gilnpdgy----------tlny 507

Qy   520 SEYLIYQPSQCRLRYLLEV 538
            :|| | ||| :|:|:||:|
Db   508 neyivypnpqvrmryllkv 526

RESULT 9
AAU29023
ID   AAU29023 standard; Protein; 534 AA.
XX
AC   AAU29023;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human PARP-2 protein #2.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia; diabetes.
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45684.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 16; Page 125-127; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia) , inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cdNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence   534 AA;

Query Match   25.5%;   Score 726.5;   DB 22;   Length 534;
Best Local Similarity 35.1%;   Pred. No. 1.8e-57;

Matches 196; Conservative 99; Mismatches 187; Indels 77; Gaps 21;

```
Qy   20 GPEKKGRQAGREED-PFRSTA----EALKAIPAEKRIIRVDPTCPLSSNPG-TQVY-- 70
           |  |::| |  || |:::||    :|::|| | ||| || ::  ||
Db    5 ggkankdrtedkqdgmpgrswaskrvsesvkallikgk-apvdpec--takvgkahvyce 61

Qy   71 --EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLE 126
          :: |||| ||||: ||||:|||||:| || | | |:|||||::|| |   -
Db   62 gndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmqqhslvacsgnln 121

Qy  127 DAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVKVDRGPVRT 182
          |  :: | |   |   ||  || ::|| :|  :   |||  | : :|
Db  122 kakeifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkp 181

Qy  183 VTKRVQPCSLDPATOKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEA 242
          | | ||:|  |  :  || :: |:  | :||||  ||:||| ||| |:|  :
Db  182 esq------ldlrvgeliklicnvqameemmemkyntkkaplgkltvaqikagyqslkk 235

Qy  243 LEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELA 302
          :: ::| |||  | ::| |:|  || |:|: || |:| |:|||  ||:||:|| |
Db  236 iedcira-gqhgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdieia 294

Qy  303 QALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHR 358
          | |  :| ::: | |  |||:| :|| || : | :| || : ::|| :::|| |   |
Db  295 ikl------vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdy 348

Qy  359 CPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTMNAVVAAILTSGLRI---SGG 414
          |||| | :|:|  ::||| | ||||  ||| : || |:| || |||| |   |:|
Db  349 tmtlldlfevekdgekeafr--edlhnrmllwhgsrmsnwvgilshglriahpeapitgy 406

Qy  415 RVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL------- 466
          ||||||||| |||||||| || :|   ||||: :|:  ||  |  |  |
Db  407 mfgkgiyfadmssksanycfasr--lkntglllsevalgqcnelleanpkaegllqgkh 464

Qy  467 ------KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQ 519
                |:| | |||:||:|  | ||| |::|| ||| | |:| || :  |::|
Db  465 stkglgkmapssahfvtlngstvplgpasdt--------gilnpdgy---------tlny 507

Qy  520 SEYLIYQESQCRLRYLLEV 538
          ||::|:|::| ||||||:|
Db  508 neyivynpnqvrmryllkv 526
```

RESULT 10
AAU29020

ID   AAU29020 standard; Protein; 534 AA.
XX
AC   AAU29020;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human PARP-2 protein #1.
XX
KW   Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW   cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;
KW   immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW   oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW   meningitis-associated intracranial complication; ischaemia;
KW   inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200164955-A1.
XX
PD   07-SEP-2001.
XX
PF   01-MAR-2001; 2001WO-US06572.
XX
PR   02-MAR-2000; 2000US-0517467.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Popoff I, Cowsert LM;
XX
DR   WPI; 2001-602570/68.
DR   N-PSDB; AAS45586.
XX
PT   Antisense compound useful for treating hyperproliferative,
PT   neurological, inflammatory and autoimmune disorders and diabetes
PT   inhibits human PARP -
XX
PS   Example 13: Page 102-104; 168pp; English.
XX
CC   The invention relates to antisense oligonucleotides targeted to human
CC   PARP nucleic acid and inhibiting expression of human PARP. PARP
CC   (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC   decondensation, DNA replication, DNA repair, gene expression, malignant
CC   transformation, cellular differentiation and apoptosis. The antisense
CC   oligonucleotide inhibitors are useful for inhibiting the expression of
CC   PARP in human cells or tissues. They are also useful for treating a
CC   human with a disease associated with PARP especially hyperproliferative
CC   disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC   neurological (e.g parkinsonism, meningitis-associated intracranial
CC   complications and ischaemia), inflammatory and autoimmune disorders (e.g
CC   arthritis) and diabetes. The present sequence is a PARP protein,
CC   the cDNA encoding which was used to design the antisense
CC   oligonucleotides.
XX
SQ   Sequence   534 AA;

Query Match   25.4%; Score 725.5; DB 22; Length 534;
Best Local Similarity   35.1%; Pred. No. 2.2e-57;
Matches 196; Conservative 99; Mismatches 187; Indels 77; Gaps 21;

```
RESULT 11
AAY51174
ID   AAY51174 standard; Protein; 570 AA.
XX
AC   AAY51174;
XX
DT   31-MAR-2000  (first entry)
XX
DE   Human brain PARP2 protein.
XX
KW   PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW   diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW   ischemic tissue damage.
XX
OS   Homo sapiens.
XX
PN   WO9964572-A2.
XX
PD   16-DEC-1999.
XX
PF   04-JUN-1999;   99WO-EP03889.
XX
PR   05-JUN-1998;   98DE-1025213.
PR   01-MAR-1999;   99DE-1008837.
XX
PA   (BADI ) BASF AG.
XX
PI   Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR   WPI; 2000-087218/07.
DR   N-PSDB; AA244287.
XX
PT   Novel genes and proteins, antibodies and binding partners useful in
PT   diagnosis and therapy of energy deficiency associated disease
PT   conditions -
XX
PS   Claim 4; Page 52-54; 96pp; German.
XX
CC   This invention describes novel human and murine poly(ADP-ribose)
CC   polymerase (PARP) homologues, which are characterised by an amino acid
CC   sequence with a functional NAD^+-binding site and no zinc finger
CC   sequence motif, of general formula CX_2CX_2MHX_2C (I). The nucleic acid
CC   sequences, PARP homologues and antibodies are useful for analytic
CC   detection of PARP homologues and for identifying PARP effectors or
CC   binding partners, as well as for determining their effectiveness.
CC   PARP-binding partners are useful for the diagnosis or therapy of a
CC   disease condition, which is the result of a PARP protein, especially an
CC   energy deficiency, which may comprise tissue damage from cell death
CC   following necrosis or apoptosis. The disease condition may be chosen
CC   from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC   in particular neurotoxic disturbances, etc. This sequence represents the
CC   human PARP2 protein used in the method of the invention.
XX
SQ   Sequence  570 AA;

Query Match             25.4%;   Score 725;  DB 21;   Length 570;
Best Local Similarity   35.4%;   Pred. No. 2.7e-57;
Matches  197;   Conservative  95;   Mismatches 185;   Indels  80;   Gaps  21;

Qy          22 EKKKGRQAG------REEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPG-TQVY---- 70
               ||| |||       :||:|| | :|:||| ::  ::| : || |
Db          46 eskkmpvaggkankdrted--kqdesvkalllkgk-apvdpec--takvgkahvycegn 99

Qy          71 EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDA 128
               :|: ||:|| ||||||:||:|||:|::| ||:||||| : ||:| |  :||||||||| :
Db         100 dvydvmlnqtnlqfnnnkyyllqlleddaqrnfsvwmrwgrvgkmqqhslvacsgnlnka 159

Qy         129 KKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVKVDRGPVRTVT 184
               |:|:||| :||||| :|: ||||| ||| ||  : |::  ::  ||:|  : |: :  |
Db         160 keifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkees1ksplkpes 219

Qy         185 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKSQQIARGFEALEALE 244
```

```
RESULT 12
AAB11480
ID   AAB11480 standard; Protein; 570 AA.
XX
AC   AAB11480;
XX
DT   02-MAR-2001  (first entry)
XX
DE   Human brain poly-ADP-ribose-polymerase protein.
XX
KW   Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
KW   cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
KW   anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
KW   antiinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
KW   2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
KW   neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
KW   Parkinson's disease; ischemic damage; microinfarction; sepsis;
KW   diabetes mellitus.
XX
OS   Homo sapiens.
XX
PN   DE19921567-A1.
XX
PD   16-NOV-2000.
XX
PF   11-MAY-1999;   99DE-1021567.
PR   11-MAY-1999;   99DE-1021567.
XX
PA   (BADI ) BASF AG.
XX
PI   Lubisch W, Sadowski J, Kock M, Hoeger T;
XX
DR   WPI; 2001-032983/05.
DR   N-PSDB; AAC82090.
XX
PT   Drugs for inhibiting PARP or especially homologous enzymes comprising
PT   4-substituted phthalazinone derivatives, useful e.g. for treating
PT   neurodegenerative disease, ischemic damage, tumors or diabetes -
XX
PS   Example A; Page 12-13; 14pp; German.
XX
CC   This invention describes novel 4-substituted 2H-phthalazin-1-one
CC   derivatives (I) which are used for the treatment or prophylaxis of
CC   diseases associated with elevated poly(ADP-ribose)-polymerase (PARP;
```

```
Db         220 q------ldlrvqeliklicnvqameemmemkyntkkaplgkltvaqikagyqslkkie 273
                       :|: |||||   | |||||| ::|:|  :||| | ||::|: :|
Qy         245 EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLVLADIELAQA 304
                       :||: | ||| ||||| | ||:|:| ::| ::| |:|| |:|
Db         274 dcira-qghgralmeacnefytriphdfglrtpplirtqkelsekiqllealgdielaik 332
               :| |:|:| ||| :| | : | :| | :|  :| | | ::|:|| ||| | || | ||:|
Qy         305 LQAVSEQEKTVEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHRCP 360
               |    | :|: |  | |:|| :| |  : |||| | | :: ||| |  :| |
Db         333 l-----vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytm 386
                      :|||: :|| ::| | | |  :|  ||  || | |: | ||: | ::|
Qy         361 TLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRV 416
               |||||:|  |   :| :| ::|| | |:|: :||| | | | |: ||:|||
Db         387 tldlfevekdgekeafr--edlhnrmllwhgsrmsnwvgilshglriappeapitgymf 444
                 | | | :| |::| | | :||| : | || | |:|| |:  |::|  | |:
Qy         417 GKGIYFASENSKSAGYVIGMKCCGAHHVGYMFLGEVALGREHHINTDNPSL-------- 466
               |||||| :||| | ::|| || | |  :|||| || |
Db         445 gkgiyfadmsskssanycfasr--lkntglllsevalggcnelleanpkaeglgqkhst 502
                          ::|||||:  :|| :|:|| |:|| | |:|| :| | ::|:| :| :|
Qy         467 ----KSPPPGFDSVIARGHTEP-DPTQDTELEDLGOQVVVPQGOPVPCPEFSSSTFSQSE 521
               :||| | | |:  |||  |:|| | ::|||| || |         || |
Db         503 kglgkmapssahfvtlngstvplgpasdt------gilnpdgy---------tlnyne 545
                   ||||||:| |:| |||||
Qy         522 YLIYQESQCRLRYLLEV 538
               :|: ||| :|:|||||:|
Db         546 yivynpnqvrmyllkv 562
```

CC   EC 2.4.2.30) activity. The products of the invention have nootropic,
CC   neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,
CC   cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,
CC   immunosuppressive, antinflammatory, antirheumatic, antiarthritic,
CC   antidiabetic. (I) are especially used for treating or preventing
CC   neurodegenerative disease or neuronal damage (specifically associated
CC   with ischemia, trauma or massive bleeding, especially apoplexy or
CC   spinal-cranial trauma), treating acute myocardial ischemia and damage
CC   Parkinson's disease), treating or preventing ischemic damage
CC   (specifically renal damage after renal ischemia or during and after
CC   kidney transplantation or heart damage after cardiac ischemia (e.g.
CC   epilepsy, specifically generalized epileptic attacks (e.g. petit mal and
CC   tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lope
CC   and complex partial attacks), treating microinfarction (e.g. during and
CC   after heart valve replacement, aneurysm resectioning and heart
CC   transplantation), revascularization of critically constricted coronary
CC   arteries (e.g. after PCTA or by-pass operations) or peripheral arteries
CC   (e.g. leg arteries), treating acute myocardial ischemia and damage during
CC   or after its mechanical or drug-induced lysis and treating tumors and
CC   their metastasis, sepsis and septic shock, inflammatory and rheumatic
CC   disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)
CC   inhibit PARP (i.e. the known form designated PARP1), they especially
CC   selectively and strongly inhibit PARP homologs, specifically the homolog
CC   PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.
CC   with K_i values of 1-20 nM) and high selectivity for PARP2 relative to
CC   PARP1 (generally by a factor of more than 5).
XX
SQ   Sequence   570 AA;

Query Match   25.4%;   Score 725;   DB 22;   Length 570;
Best Local Similarity 35.4%;   Pred. No. 2.7e-57;
Matches 197;   Conservative 95; Mismatches 185;   Indels 80;   Gaps   21;

Qy   22 EKKKGRQAG------REEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPG-TQVY---- 70
            ::||| || |::||| |:| || ::|| | || | || : ||
Db   46 eskkmpvaggkankdrted--kqdesvkalllkgk-apvdpec--takvgkahvycegn 99

Qy   71 EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDA 128
            : |: |||||: |||||: ||| | | | :|:| | | | | |: || | ::|:|
Db   100 dvydvmlngtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnka 159

Qy   129 KKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEAQEAVVKVDRGPVRTVT 184
            |: |||| :|||||| ||| || :| |||| ::| : |:: |: :||: |
Db   160 keifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpes 219

Qy   185 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKMKPLGKLSKQQIARGFEALEALE 244
            | || |:|:| : || | : | | :| | | |:| :||:|:| ||:|
Db   220 q------ldlrvgeliklicnvqameemmmenkyntkkaplglktvaqikagyyqslkkie 273

Qy   245 EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQA 304
            || :|| :|:| |||| || ||| | | | : | || |: | |: |||:|
Db   274 dcira-gqhgralmeacnefytriphdfglrtpllirtqkelsekiqllealgdieialk 332

Qy   305 LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHRCP 360
            | :| | || | |: | | |||| ||::| | : |:| ||:| | |
Db   333 l------vktelgspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdytm 386

Qy   361 TLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRV 416
            | | :: | || | : || | |||| | | || |
Db   387 tldlfevedgekeafr--edlnrmllwhgsrmsnwvgilsnglrriappeapitgymf 444

Qy   417 GKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL--------- 466
            ||| |||| || ||| | | || | | ||| | |:: | | |
Db   445 gkglyfadmssksanycfasr--lkntgllllsevalggcnelleanpkaegllqqkhst 502

Qy   467 ---KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSE 521
               | | | | |: ||| || || | :|
Db   503 kglgkmapssahfvtlngstvplgpasdt-------gilnpdgy----------tlnyne 545

Qy   522 YLIYQESQCRLRYLLEV 538
            ::| :| | |||
Db   546 yivvpnqvrmryllkv 562

RESULT 13
AAB60693
ID   AAB60693 standard; Protein; 521 AA.
XX
AC   AAB60693;
XX
DT   11-MAY-2001 (first entry)
XX
DE   Human poly(ADP-ribose) synthetase sbhPARS2.
XX
KW   Human; poly(ADP-ribose) synthetase; sbhPARS2; vaccine; drug screening;
KW   ischaemic disorder; cerebral ischaemia; cardiac ischaemia;
KW   myocardial infarction; stroke; inflammation; autoimmune disease;
KW   diabetes; multiple sclerosis; neurodegenerative disease;
KW   Parkinson's disease; Alzheimer's disease; chromosome localisation.
XX
OS   Homo sapiens.
XX
PN   WO200112645-A1.
XX
PD   22-FEB-2001.
XX
PF   10-AUG-2000; 2000WO-US21775.
XX
PR   12-AUG-1999; 99US-0373441.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI   Barone F, Field J, Kabnick KS, Li X, McQueney MS, Zhu Y;
XX
DR   WPI; 2001-211196/21.
DR   N-PSDB; AAF59996.
XX
PT   New human polypeptide of the polyADPribose synthetase family for use in
PT   diagnostic assays and for screening modulators used for preventing and
PT   treating inflammation, autoimmune disease and Alzheimers disease -
XX
PS   Claim 1; Page 21; 30pp; English.
XX
CC   This sequence represents a novel human poly(ADP-ribose)
CC   synthetase, sbhPARS2. The invention also relates to fragments,
CC   variants and sequences with at least 95% identity to the sbhPARS2
CC   protein or nucleotide sequence; expression systems and host cells
CC   comprising an sbhPARS2 nucleic acid sequence; the recombinant expression
CC   of sbhPARS2; and an antibody specific for sbhPARS2. sbhPARS2 proteins
CC   and nucleotides are useful as vaccines for inducing an immunological
CC   response in a mammal. The sbhPARS2 protein is useful for identifying
CC   compounds which inhibit or stimulate its activity or expression level.
CC   Such agonists and antagonists of sbhPARS2 are useful for treating human
CC   diseaes including ischaemia and ischaemic tissue injury (e.g., cerebral
CC   and cardiac ischaemia, myocardial infarction, stroke), inflammation,
CC   autoimmune disease (e.g. diabetes, multiple sclerosis) and
CC   neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's
CC   disease). sbhPARS2 nucleic acids are useful as diagnostic reagents for
CC   detecting mutations in the associated gene; as hybridisation probes
CC   to isolate full-length sbhPARS2 cDNAs and sbhPARS2 genomic clones; and
CC   for chromosome localisation studies. The sbhPARS2 protein is also useful
CC   as an immunogen to produce antibodies for therapeutic use. sbhPARS2
CC   proteins, nucleotides and antibodies are also useful in screening methods
CC   for detecting the effect of added compounds on the production of mRNA and
CC   protein in cells.
XX
SQ   Sequence   521 AA;

Query Match   25.3%;   Score 722;   DB 22;   Length 521;
Best Local Similarity 34.7%;   Pred. No. 4.4e-57;
Matches 192;   Conservative 99; Mismatches 184;   Indels 78;   Gaps   20;

Qy   20 GPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPG-TQVY----EDYN 74
            |:: || :: || |:|||:| || ||| |: | || : | ::

```
DR    N-PSDB; AAZ60617.
XX
PT    Modulating cell death, growth and stress resistance in eukaryotes,
PT    specifically plants, used, e.g. to impart fungus or nematode resistance
PT    .
XX
PS    Claim 18; Page 99-101; 126pp; English.
XX
CC    The present sequence represents a NAP protein. This protein is a
CC    poly(ADP-ribose) polymerase (PARP) protein (also known as
CC    poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
CC    cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide
CC    sequences can be used for modulation of programmed cell death in
CC    eukaryotic cells. The method is used, specifically in plants, to induce,
CC    or protect against, programmed cell death, depending on the extent to
CC    which PARP activity is reduced. Reducing expression of endogenous NAP
CC    class PARP only is also used to modulate programmed cell death, to
CC    increase growth rate and to produce plant cells that are more tolerant
CC    of stress (cold, chemical treatments, pathogens, pests, drought, heat,
CC    etc., or during transformation). Particular applications are generation
CC    of plants that are resistant to fungi or nematodes; are male or female
CC    sterile; or have better seed-shatter properties. The methods are also
CC    used to improve growth of transformed plant cells (and derived calli or
CC    complete plants).
XX
SQ    Sequence    637 AA;

      Query Match              24.1%;    Score 688.5;    DB 21;    Length 637;
      Best Local Similarity    36.2%;    Pred. No. 7.2e-54;
      Matches  174;    Conservative  89;    Mismatches  185;    Indels  33;    Gaps  16;

Qy     73  YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK 130
               ::||||| :|||||  | |:|||::: | ||:|
Db    169  ydailngtnvrdnnkffvlqvlesdskktymvytrwgrvgvkggskldgpydswdraie 228

Qy    131  DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQP- 189
               :| :||||:|| ||: | :| | :| : |:  ::|
Db    229  iftnkfndktknywsdrkefiphpksytwlemdygkeendspvnnd--ipssssevkpe 285

Qy    190  -CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 248
               || || | : :: | || || :| :|: :|| :||::| ::|
Db    286  qskldtrvakfislicnvsmmaqhmmeignynanklplgkisksstiskgyevlkrisevi- 344

Qy    249  GPTDGGQSLEELSSHFYTVIPHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ 306
               ||||||||||| : | :|  ||||||| :|: :|  | :|||  | || |:|||||||||
Db    345  -drydrtrleelsgefytviphdfgfkmsqfvidtpgklkqkiemvealgelelatkil 403

Qy    307  AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ- 363
               | :| ||  ||||| |  :  | ||: | |:| || | :| |: |:| |:|
Db    404  svdpglq-----ddplyyhyqqlncgltpvgndseefsmvanymenthakthsgytveia 458

Qy    364  HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGKG 419
               || ||||| : || | |:| |::::|||| |||||| ||||:||| |: :||||:||
Db    459  qifrasraveadrfqqfsssknrmllwhgsrltnwagilsqglriappeapvtgymfgkg 518

Qy    420  IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 479
               :|||| || || |||||| | | ||| |:|||| || |::|| ||| :|:|
Db    519  vyfadmfsksanycya-ntgand-gvillcevalgdmnellysdynadnlppgklstkgv 576

Qy    480  GHTEPDPTQDTELELDGQQVVVPQGQPV--PCPEFSSSTFSQSEYLIYQESQCRLRYLLE 537
               | | ||  |: : ||| | : |||  : ::| |::|||| :||||||||:|
Db    577  gktapnpseaqtle-dg---vvvplgkpversc---skgmllyneyivynveqikmryyiq 630

Qy    538  V 538
          |
Db    631  v 631
```

```
RESULT 14

AAY68835
ID    AAY68835 standard; Protein; 637 AA.
XX
AC    AAY68835;
XX
DT    16-MAY-2000  (first entry)
XX
DE    The poly(ADP-ribose) polymerase NAP protein of Arabidopsis.
XX
KW    NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
KW    programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
KW    pest; drought; heat; fungi; nematode; seed-shatter.
XX
OS    Arabidopsis thaliana.
XX
FH    Key             Location/Qualifiers
FT    Misc-difference 1..138
FT                    /note= "these residues are specifically claimed in
FT                    claim 18"
XX
PN    WO200004173-A1.
XX
PD    27-JAN-2000.
XX
PF    12-JUL-1999;  99WO-EP04940.
XX
PR    17-JUL-1998;  98US-0118276.
XX
PA    (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX
PI    Bablychuk E, Kushnir S, De Block M;
XX
DR    WPI; 2000-182436/16.
```

```
Db      5  ggkankdrtedkqd-------esvkalllkgk-apvdpec--takvgkahvcegndvyd 54
               :|| ||||| ||||:|| |||||:|| |||||:  ||: ||
Qy     75  CTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SKINHFTRLEDAKKDF 132
               :||| |:|||||| :||||||:::  |  |:||::  || :|
Db     55  vmlngtnlqfnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhslvacsgnlnkakeif 114

Qy    133  EKKFREKTKNNWAERDHFVSHPGKYTLIEV----QAEDEAQEAVVKVDRGPVRTVTKRVQ 188
               :|| :||||||:|| |: |:: ||| | : ||  : ||: :|
Db    115  qkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkeeslksplkpesq--- 171

Qy    189  PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 248
               | :||:|:: :| |:|:|: || |||::|||:|:|:
Db    172  ---ldlrvgelikliicnvqaneemnunemkyntkkaplgkltvaqikagyqsikkiedcir 228

Qy    249  GPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAV 308
               |:|||||||| :|||||||||| | ::| :| || |:|||  | |:|:|:|||||
Db    229  a-gqhgralmeacnefytriphdfgjirtpplirtqkelsekiqlealgdlelaikl--- 284

Qy    309  SEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS---NHRCPTLQH 364
               || |:|| || :|:|:: :  |:  : || |: |
Db    285  ---vktelqspehpldqhyrnlhcalrpldhesyefkvisqyilgsthapthsdytmtlld 341

Qy    365  IWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGI 420
               :: ||| | :  :| |||||| || |:|| |:|: :
Db    342  lfevekdgekeafr--edlhnrmlwhgsrmsnwvvilshglriappeapitgymfgkgi 399

Qy    421  YFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL----------- 466
               :| ::|||| |
Db    400  yfadmssksanycfasr--lkntglillsevalgqcnelleanpkaeglligqkhstkglg 457

Qy    467  KSPPPGFDSVIARGHTEP-DPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIY 525
               |:|| |||| | |:||| ||: | :: | :||
Db    458  kmapssahfvtlngstvplgpasdt--------glinpdgy------------tlnyneyivy 500

Qy    526  QESQCRLRYLLEV 538
               ::|| ||:|||:
Db    501  npnqvtmryllkv 513
```

```
RESULT 15

AAB93513
ID    AAB93513 standard; Protein; 531 AA.
XX
```

```
AC    AAB93513;
XX
DT    26-JUN-2001  (first entry)
XX
DE    Human protein sequence SEQ ID NO:12844.
XX
KW    Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS    Homo sapiens.
XX
PN    EP1074617-A2.
XX
PD    07-FEB-2001.
XX
PF    28-JUL-2000;  2000EP-0116126.
XX
PR    29-JUL-1999;  99JP-0248036.
PR    27-AUG-1999;  99JP-0300253.
PR    11-JAN-2000;  2000JP-0118776.
PR    02-MAY-2000;  2000JP-0183767.
PR    09-JUN-2000;  2000JP-0241899.
XX
PA    (HELI-)  HELIX RES INST.
XX
PI    Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI    Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR    WPI; 2001-318749/34.
XX
PT    Primer sets for synthesizing polynucleotides, particularly the 5602
PT    full-length cDNAs defined in the specification, and for the detection
PT    and/or diagnosis of the abnormality of the proteins encoded by the
PT    full-length cDNAs -
XX
PS    Claim 8; SEQ ID 12844; 2537pp + CD ROM; English.
XX
CC    The present invention describes primer sets for synthesising 5602
CC    full-length cDNAs defined in the specification. Where a primer set
CC    comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC    to the complementary strand of a polynucleotide which comprises one of
CC    the 5602 nucleotide sequences defined in the specification, where the
CC    oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC    of an oligonucleotide comprising a sequence complementary to the
CC    complementary strand of a polynucleotide which comprises a 5'-end
CC    sequence and an oligonucleotide comprising a sequence complementary to a
CC    polynucleotide which comprises a 3'-end sequence, where the
CC    oligonucleotide comprises at least 15 nucleotides and the combination of
CC    the 5'-end sequence/3'-end sequence is selected from those defined in
CC    the specification. The primer sets can be used in antisense therapy and
CC    in gene therapy. The primers are useful for synthesising polynucleotides,
CC    particularly full-length cDNAs. The primers are also useful for the
CC    detection and/or diagnosis of the abnormality of the proteins encoded by
CC    the full-length cDNAs. The primers allow obtaining of the full-length
CC    cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC    AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC    AAH13633 to AAH18742 represent human amino acid sequences; and AAH13629 to AAH13632
CC    AAB95893 represent human amino acid sequences, all of which are used in the exemplification
CC    represent oligonucleotides, all of which are used in the exemplification
CC    of the present invention.
XX
SQ    Sequence  531 AA;

Query Match          23.9%;   Score 682.5;  DB 22;  Length 531;
Best Local Similarity   36.9%;   Pred. No. 1.9e-53;
Matches  184;  Conservative  86; Mismatches 173;  Indels  55;  Gaps  18;

Qy     22 EKKKGRQAG------REED-----PFRSTA-----EALKAIPAEKRIIRVDPTCPLSSNP 65
             || ||    |||  ||     || ||      |:||:| : ||| : :
Db     46 eskkmpvaggkankdrtedkqdgmpgrswaskrvsesvkalllkgk-apvdpec--takv 102

Qy     66 G-TQYY----EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SK 118
           | || |    |||||:||||||||||||:|||:|  : |  : | : |||||:|::|| :
Db    103 gkahvycegndvydvmlnqtnlqfnnnkyyliqlleddaqrnfsvwmrwgrvgkmgqhsl 162
```

(Qy/Db alignment, right side)

```
Qy  119 INHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEAQEAVVK 174
          :  |: |:||| :||||| :|||||| :|:|    :||  :||   :||  :
Db  163 vacsgnlnkakeifqkkfldktknnwedrekfekvpgkydmlqmdyatntqdeeetkkee 222

Qy  175 VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA 234
          : |:: :|  |:||  |:  | | ::| |||  :  | | ||||| : ||
Db  223 slksplkpesq------ldlrvqeliklicnvqameemmemkyntkkaplgkitvaqik 276

Qy  235 RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL 294
          |||||: ::|:: :  :  || |: | : |:| ||:||  : :|
Db  277 agygslkkiedcira-gqhgralmeacnefytriphdfglirtqkelsekiqlle 335

Qy  295 VLADIELAQALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 353
          | ||| |:|            |::|   :|  :|  | : | :|: |:|:  :|
Db  336 algdieiaikl-----vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqst 389

Qy  354 GS--NHRCPTLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP 410
          |:    |  :|  :||:|| : ||  |   ||  :| |  |:|: :| :| :||: :|
Db  390 hapthsdytmtlldlfevekdgekeafr--edlnrmliwhgsrmsnwygilshgiriap 447

Qy  411 H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL 466
          :|    |||||  |:|||: |||    | :  |  :| | :|: ||::  ::   :  |
Db  448 peapitgymfgkgylyfadmssksanycfasr--lkhtglllsevalgqcnelleanpka 505

Qy  467 KSPPPGFDSVIARGHTEP 484
          | ::|| |  |  |  :|
Db  506 egllqgkhstkglgkmap 523
```

Search completed: August 29, 2002, 07:57:16
Job time: 293 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on: August 29, 2002, 08:02:15 ; Search time 36.19 Seconds
(without alignments)
577.744 Million cell updates/sec

Title: US-09-701-586B-6
Perfect score: 2854
Sequence: 1 MSLLFLAMAPKPKPWVQTEG.........EYLIYQESQCRLRYLLEVHL 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2811	98.5	533	1	PPO3_HUMAN	Q9y6f1 homo sapien
2	738	25.9	559	1	PPO2_MOUSE	O88554 mus musculu
3	729.5	25.6	583	1	PPO2_HUMAN	Q9ugn5 homo sapien
4	688.5	24.1	637	1	PPOL_ARATH	Q11207 arabidopsis
5	681	23.9	1015	1	PPOL_BOVIN	P18493 bos taurus
6	674.5	23.6	996	1	PPOL_SARPE	Q11208 sarcophaga
7	674.5	23.6	1011	1	PPOL_CHICK	P26446 gallus gall
8	673.5	23.6	1012	1	PPOL_CRIGR	Q9152 cricetulus
9	670	23.5	1013	1	PPOL_HUMAN	P09874 homo sapien
10	665.5	23.3	1013	1	PPOL_RAT	P27008 rattus norv
11	653	22.9	998	1	PPOL_XENLA	P31669 xenopus lae
12	653	22.9	1012	1	PPOL_MOUSE	P11103 mus musculu
13	638	22.4	994	1	PPOL_DROME	P35875 drosophila
14	393.5	13.8	538	1	YQN4_CAEEL	Q09525 caenorhabdi
15	258	9.0	1724	1	PPOV_HUMAN	Q9ukk3 homo sapien
16	190.5	6.7	135	1	PPOL_ONCMA	O08824 oncorhynchu
17	126	4.4	1222	1	YMP3_CAEEL	Q10947 caenorhabdi
18	120	4.2	1938	1	MYS_AEQIR	P24733 aequipecten
19	115.5	4.0	1395	1	SP41_YEAST	P38904 saccharomyc
20	114	4.0	918	1	HXK1_BOVIN	P27595 bos taurus
21	113.5	4.0	1164	1	BAG_STRAG	P27951 streptococc
22	113.5	4.0	4540	1	DYHC_PARTE	Q27171 paramecium
23	109.5	3.8	446	1	GAG_OMVVS	P16900 ovine lenti
24	108.5	3.8	382	1	MYB_AVIMB	P01104 avian myelo
25	107	3.7	1186	1	SMC_BACSU	P51834 bacillus su
26	106	3.7	724	1	HMMR_HUMAN	O75330 homo sapien
27	106	3.7	932	1	YAIA_SCHPO	Q09897 schizosacch
28	106	3.7	1057	1	POL_SIVAI	Q02836 simian immu
29	105.5	3.7	2025	1	TTC3_HUMAN	P53804 homo sapien
30	105	3.7	882	1	RA50_PYRFU	P58301 pyrococcus
31	104.5	3.7	641	1	MYB_CHICK	P01103 gallus gall
32	104.5	3.7	1403	1	VG22_HSVI1	Q00105 ictalurid h
33	104	3.6	1324	1	CUT3_SCHPO	P41004 schizosacch
34	104	3.6	1433	1	REST_CHICK	O42184 gallus gall
35	103	3.6	716	1	RRP2_IALE2	P26123 influenza a
36	103	3.6	1940	1	MYH3_RAT	P12847 rattus norv
37	102.5	3.6	640	1	MYH_HUMAN	P10242 homo sapien
38	102.5	3.6	704	1	VPS1_YEAST	P21576 saccharomyc
39	102.5	3.6	3210	1	CENF_HUMAN	P49454 homo sapien
40	102	3.6	716	1	RRP2_IALE1	P26122 influenza a
41	101.5	3.6	584	1	LIGA_HUMAN	P41214 homo sapien
42	101.5	3.6	612	1	EXO2_BPT5	P11109 bacteriopha
43	101.5	3.6	874	1	SYA_PASMU	P57933 pasteurella
44	101.5	3.6	3075	1	LMA1_HUMAN	P25391 homo sapien
45	101	3.5	782	1	L100_ADE12	P36714 human adeno

ALIGNMENTS

RESULT 1
PPO3_HUMAN

ID   PPO3_HUMAN   STANDARD;   PRT;   533 AA.
AC   Q9Y6F1; Q9UG81;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+) ADP-
DE   ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (pADPRT-3)
DE   (hPARP-3).
GN   ADPRTL3 OR PARP3 OR ADPRT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=99263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [2]
RP   SEQUENCE OF 75-533 FROM N.A.
RC   TISSUE=kidney;
RA   Ansorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
-!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
     nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
-!- SUBCELLULAR LOCATION: Nuclear (Potential).
-!- TISSUE SPECIFICITY: Widely expressed; the highest levels are in
     the kidney, skeletal muscle, liver, heart and spleen; also
     detected in pancreas, lung, placenta, brain, leukocytes, colon,
     small intestine, ovary, testis, prostate and thymus.
-!- SIMILARITY: BELONGS TO THE PARP FAMILY.

DR   EMBL; AF083068; AAD29855.1; -.
DR   EMBL; AL050034; CAB43246.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation.
FT   DOMAIN   14   20   NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   CONFLICT   80   80   K -> N (IN REF. 2).

```
FT   CONFLICT     171    171       A -> G (IN REF. 2).
FT   CONFLICT     411    411       K -> E (IN REF. 2).
SQ   SEQUENCE   533 AA;  60117 MW;  7C0AB89E64D1B9FD CRC64;

Query Match             98.5%;  Score 2811;  DB 1;  Length 533;
Best Local Similarity   99.6%;  Pred. No. 1.3e-188;
Matches 531;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    8 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60

Qy   68 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWRGVGEVGQSKINHFTRLED 127
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWRGVGEVGQSKINHFTRLED 120

Qy  128 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 187
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRAPVRTVTKRV 180

Qy  188 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEAAL 247
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEAAL 240

Qy  248 KGPTDGGQSLEELLSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 307
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KGPTDGGQSLEELLSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQA 300

Qy  308 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 367
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 360

Qy  368 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKIYFASENS 427
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKIYFASENS 420

Qy  428 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 487
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480

Qy  488 QDTELELDGQQVVVPQGQPVPCPEFSSSSTFSQSEYLIYQESQCRLRYLLEVHL 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 QDTELELDGQQVVVPQGQPVPCPEFSSSSTFSQSEYLIYQESQCRLRYLLEVHL 533

RESULT   2
PPO2_MOUSE
ID   PPO2_MOUSE     STANDARD;      PRT;    559 AA.
AC   O88554; Q99N29;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE   (mPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC   TISSUE=Embryo;
RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129Sv;
```

```
RX   MEDLINE=21179160; PubMed=11133988;
RA   Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,
RA   Niedergang C.P.;
RT   "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2
RT   (PARP-2) gene to the gene for RNase P RNA.";
RL   J. Biol. Chem. 276:11092-11099(2001).
RN   [3]
RP   SEQUENCE OF 9-559 FROM N.A.
RC   STRAIN=129/Sv X C57BL/6;
RX   MEDLINE=99268466; PubMed=10338144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC       SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in
CC       testis followed by ovary.
CC   -!- INDUCTION: By high levels of DNA-damaging agents.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; AJ007780; CAA07679.1; -.
DR   EMBL; AF191547; AAK13253.1; -.
DR   EMBL; AF072521; AAC25415.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1341112; Adprt2.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND      1     65       POTENTIAL.
FT   DOMAIN       66    559       NAD-BINDING (BY SIMILARITY).
FT   DOMAIN        3      9       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                (POTENTIAL).
FT   DOMAIN       33     39       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                (POTENTIAL).
FT   CONFLICT     82     82       L -> V (IN REF. 2).
FT   CONFLICT    177    177       V -> I (IN REF. 2).
FT   CONFLICT    486    486       R -> Q (IN REF. 2).
SQ   SEQUENCE   559 AA;  63396 MW;  E0AEDAEE412C1445 CRC64;

Query Match             25.9%;  Score 738;  DB 1;  Length 559;
Best Local Similarity   34.3%;  Pred. No. 4.4e-44;
Matches 187;  Conservative 103;  Mismatches 191;  Indels 64;  Gaps 19;

Qy   23 KKKGRQA-GREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVY----EDYNCTL 77
          ::||.||| :::|: ||||||||||||||||||: :.|||: : |:|| :| ::|
Db   42 QRKGPMAGGKDADRTKDNRDSVKTLLLKGK-APVDPECAAKLGK-AHVYCEGDDVYDVML 99

Qy   78 NQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWRGVEGVGQ-SKINHFTRLEDAKKDFEKK 135
          |||| :||||.||||:||: |::|: :|||.|.||.|.|| :|..|| |:||::|:||
Db  100 NQTNLQFNNNKYYLIQLLEDDAQRNFSVWNRWGRVGKTGQHSLVTCSGDLNKAKEIFQKK 159

Qy  136 FREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEA---QEAVVKVDRGPVRTVTKRVQ 188
          |::|||||| :|:|: | ||||   :|     ::|:      ::|: :| :|:
Db  160 FLDKTKNNWEDRENFEKVPGKYDMLQMDYAASTQDESKTKEEETLKPE----------- 207

Qy  189 PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 248
```

```
Db   208  -SQLDLRVQELLKLICNVQTMEEMMIEMKYDTKRAPLGKLTVAQIKAGYQSLKKIEDCIR 266
            ||      |::    |:  : |      ||:: ||| :::|| :: : : ::
Qy   249  GPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQAV 308
           : | |: ||| ||| :||: |  : || :  |  || :  ||| : |
Db   267  A-GQHGRALVEACNEFYTRIPHDFGLSIPPVIRTEKELSDKVKLLEALGDIEI--ALKLV 323
            : ||: |  :   |   |  || :|| ||    |   || : :|
Qy   309  SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHR--CPTLQHI 365
           ||||:    |::||   ||  |:|    : |:|   | :  |:|      |    |: |||
Db   324  KSERQGLE---HPLDQHYRNLHCALRPLDHESNEFKVISQYLQSTHAPTHKDYTMTLLDV 380
            || |  |        || :||  |       |  |   || : :| | |     ||||
Qy   366  WKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVKGIY 421
            :: ||| :: |:    || ||||| :  : ||| ||:|
Db   381  FEVEKEGEKEAFR--EDLPNRMLLWHGSRLSNWVGILSHGLRVAPPEAPITGYMFGKGIY 438
           ||||    || |     |  : |  |    |:     |  | : | :|||| |  ||| |
Qy   422  FASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGH 481
            |  :|||||    |   ||  :|: :  |  | | |:   |   | || | :  ||||
Db   439  FADMSSKSANYCFASR--LKNTGLLLLSEVALGQCNELLEANPKAQGLLRGKHSTKGMGK 496
            |  :|| |  |         |  |  ||| :   :  |  |  || :||    | |||
Qy   482  TEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSS-------TFSQSEYLIYQESQCRLR 533
             |         | |              | |             |::|| :| :||:|
Db   497  MAPSPA---------HFITLNGSTVPLGPASDTGILNPEGYTLNYNEFIVISPNQVRMR 546
                    |||:|   |||::|
Qy   534  YLLEV 538
           |||::
Db   547  YLLKI 551
```

RESULT 3
PPO2_HUMAN
ID   PPO2_HUMAN   STANDARD;   PRT;   583 AA.
AC   Q9UGN5; Q9Y6C8; Q9NUV2; Q9QMR4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE   ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE   (hPARP-2).
GN   ADPRTL2 OR PARP2 OR ADPRT2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99292755; PubMed=10364231;
RA   Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,
RA   Muller S., Hoeger T., Menissier-de Murcia J., de Murcia G.M.;
RT   "PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose)
RT   polymerase.";
RL   J. Biol. Chem. 274:17860-17868(1999).
RN   [2]
RP   SEQUENCE OF 2-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal brain;
RX   MEDLINE=99263509; PubMed=10329013;
RA   Johansson M.;
RT   "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA
RT   cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL   Genomics 57:442-445(1999).
RN   [3]
RP   SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).
RC   TISSUE=Fibroblast;
RX   MEDLINE=99268466; PubMed=10338144;
RA   Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT   "pADPRT-2: a novel mammalian polymerizing(ADP-ribosyl)transferase gene
RT   related to truncated pADPRT homologues in plants and Caenorhabditis
RT   elegans.";
RL   FEBS Lett. 449:259-263(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA   Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA   Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA   Masuo Y., Kanehori K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.
CC   SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (BY
CC   similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC   produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN
CC   THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO
CC   DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;
CC   LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND
CC   THYMUS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AJ236912; CAB65088.1; -.
DR   EMBL; AF085734; AAD29857.1; ALT_INIT.
DR   EMBL; AJ236876; CAB41505.2; ALT_INIT.
DR   EMBL; AK001980; BAA92017.1; ALT_TERM.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Alternative splicing.
FT   DNA_BIND   1   88   POTENTIAL.
FT   DOMAIN   86   583   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   4   7   NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT   (POTENTIAL).
FT   DOMAIN   35   40   NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT   (POTENTIAL).
FT   VARSPLIC   68   80   MISSING (IN ISOFORM 2).
FT   CONFLICT   447   447   P -> H (IN REF. 2).
FT   CONFLICT   481   481   N -> H (IN REF. 4).
SQ   SEQUENCE   583 AA; 66205 MW; 5B7AE8AE531836AF CRC64;

Query Match   25.6%; Score 729.5; DB 1; Length 583;
Best Local Similarity   35.4%;   Pred. No. 1.8e-43;
Matches 201; Conservative 95; Mismatches 184; Indels 87; Gaps 22;

Qy   22   EKKKGRQAG------REED------PFRSTA------EALKAIPAEKRIIRVDPTCPLSSNP 65
           :|| |::| | :||| | | | || : |: |||
Db   46   ESKKMPVAGGKANKDRTEDKQDGMPGRSWASKRVSESVKALLLGK-APVDPEC--TAKV 102
Qy   66   G-TQVY-----EDYNCTLNQTNIENNNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQ-SK 118
           | || : | : |||||| || | |||::| : | | | ||:||| | :||
Db   103 GKAHVYCEGNDVYDVMLNQTNLQFNNNKYYLIQLLEDDAQRNFSVWMRWGRVGKMGQHSL 162
Qy   119 INHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEV---QAEDEAQEAAVVK 174
           | | |||||||| ||| |||| | | ||| || ||||||| || | : |:|
Db   163 VACSGNLNKAKEIFQKFLDKTKNWEDREKFEKVPGKYDMLQMDYATNTQDEEETKKEE 222
Qy   175 VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA 234
           | ||| : |||: | | | ||| : || | | :| | |:|| | || | |
Db   223 SLKSPLKPESQ------LDLRVQELIKLICNVQAMEEMMEMKYNTKKAPLGKLTVAQIK 276
Qy   235 RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL 294
```

```
Db   277 ...AGYQSLKKIEDCIRA-GQHGRALMEACNEFYTRIPHDFGLRTPPLLRTQKELSEKIQLLE 335
Qy   295 VLADIELAQALQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 353
Db   336 ALGDIEIAIKL------VKTELQSPEHPLDQHYRNLHCALRPLDHESYEFKVISQYLQST 389
Qy   354 GS--NHRCPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP 410
Db   390 HAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGILSHGLRIAP 447
Qy   411 H---SGGRVGKGITFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSL 466
Db   448 PEAPTTGYMFGKGIYFADMSSKSANYCFASR--LKNTGLLLLSEVALGQCNELLEANPKA 505
Qy   467 ------KSPPPGFDSVIARGHTEP-DPTQDTELEDGQQVVVPQGQPVPCPE 511
Db   506 EGLLQGKHSTKGLGKMAPSSAHFVTLNGSTVPLGPASDT------GILNPDGY------ 552
Qy   512 FSSSTFSQSEYLIYQESQCRLRYLLEV 538
Db   553 ----TLNYNEYIVYNPNQVRMRYLLKV 575
```

RESULT 4
PPOL_ARATH
ID   PPOL_ARATH   STANDARD;   PRT;   637 AA.
AC   Q11207;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   APP.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95269779; PubMed=7750552;
RA   Lepiniec L., Babiychuk E., Kushnir S., van Montagu M., Inze D.;
RT   "Characterization of an Arabidopsis thaliana cDNA homologue to animal
RT   poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 364:103-108(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the European Bioinformatics
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; Z48243; CAA88288.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.

DR   SMART; SM00513; SAP; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation.
FT   DNA_BIND   1   140   POTENTIAL.
FT   DOMAIN   1   140   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   41   62   NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   637 AA; 72175 MW; 527A8F46460SD127 CRC64;

Query Match   24.1%; Score 688.5; DB 1; Length 637;
Best Local Similarity   36.2%; Pred. No. 1.5e-40;
Matches 174; Conservative   89; Mismatches 185; Indels   33; Gaps 16;

```
Qy    73 YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK 130
Db   169 YDAILNQTNVRDNNNKFFVLQVLESDSKKTYMYTRNRGRVGKGQSKLDGPYDSWDRAIE 228
Qy   131 DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRRTVTKRVQP- 189
Db   229 IFTNKFNDKTKNYWSDRKEEFIPHPKSYTWLEMDYGKEENDSPVNND---IPSSSSEVKPE 285
Qy   190 -CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 248
Db   286 QSKLDTRVAKFISLICNVSMAQHMMEIGYNANKLPLGKISKSTISKGYEVLKRISEVI- 344
Qy   249 GPTDGGQSLEELSSHFYTVIPHFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ 306
Db   345 -DRYDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLL 403
Qy   307 AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ- 363
Db   404 SVDPGLQ-----DDPLYYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVEIA 458
Qy   364 HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGKG 419
Db   459 QLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKG 518
Qy   420 IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 479
Db   519 VYFADMFSKSANYCYA-NTGAND-GVLLLCEVALGDMNELLYSDYNADNLPPGKLSTKGV 576
Qy   480 GHTEPDPTQDTELEDGQQVVVPQGQPV--PCPEFSSSTFSQSEYLIYQESQCRLRYLLE 537
Db   577 GKTAPNPSEAQTLE-DG--VVVPLGKPVERSC--SKGMLLYNEYIVYNVEQIKMRVVIQ 630
Qy   538 V 538
Db   631 V 631
```

RESULT 5
PPOL_BOVIN
ID   PPOL_BOVIN   STANDARD;   PRT;   1015 AA.
AC   P18493; Q9TS00;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90382673; PubMed=2119324;
RA   Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;
RT   "Cloning of a full-length cDNA encoding bovine thymus
RT   poly(ADP-ribose) synthetase: evolutionarily conserved segments and
RT   their potential functions.";
RL   Gene 90:249-254(1990).

```
RN   [2]
RP   SEQUENCE OF 647-714 AND 838-903 FROM N.A.
RX   MEDLINE=88151954; PubMed=2450019;
RA   Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,
RA   Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;
RT   "Depression in gene expression for poly(ADP-ribose) synthetase during
RT   the interferon-gamma-induced activation process of murine macrophage
RT   tumor cells.";
RL   Eur. J. Biochem. 171:571-575(1988).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D90073; BAA14114.1; -.
DR   EMBL; X06986; CAA30046.1; -.
DR   EMBL; X06987; CAA30047.1; -.
DR   PIR; JS0428; JS0428.
DR   PIR; S00328; S00328.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS00347; PARP_ZN_FINGER_2; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0
FT   DNA_BIND      1    374                   BY SIMILARITY.
FT   DOMAIN      375    525                   AUTOMODIFICATION DOMAIN.
FT   DOMAIN      386    462                   BRCT.
FT   DOMAIN      526   1015                   NAD-BINDING.
FT   ZN_FING     127     55                   PARP-TYPE.
FT   ZN_FING      20    164                   PARP-TYPE.
FT   DOMAIN      209    211                   NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      223    228                   NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     408    408                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     414    414                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     446    446                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     449    449                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     457    457                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     472    472                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     489    489                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     492    492                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     514    514                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     515    515                   ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     521    521                   ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE   1015 AA;  113355 MW;  0A5FE9D9F04F5B04 CRC64;

Query Match             23.98%;  Score 681;  DB 1;  Length 1015;
Best Local Similarity   33.9%;
Matches  188;  Conservative  99;  Mismatches  205;  Indels  62;  Gaps  24;

Qy      7 AMAPKPK----PWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLS 62
            |::|||     :::||::::||   :
Db    493 AVGPKGKSGAAPSKKSKGPVKEEG-TNKSEKRMKLTLKGGAA-------VDPDSGLE 541

Qy     63 SNPGTQYYED----YNCTLNQTNIENNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV-GQ 116
            -|--|| ::||    :::|||| :: |:||::  || | |||
Db    542 HN--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVIGS 599

Qy    117 SKINHFTRLEDAKKDFEKKFREKTKNNWAEERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKV 175
            || ||   :||| |:| |::::: ||:: || :: |:| | |:: || :: |||| |
Db    600 NKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKHPKKFYPLEIDYGQDE--EAVKKL 656

Qy    176 DRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKILSKQQIAR 235
            :||| | |::|||::  |::|||:||:|: ||:|| ||||:|:|||:|| ||:||:||||
Db    657 TVNP-GTKSKLPKP------VQNLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQA 709

Qy    236 GFEALEALEEAL-KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL 294
            | |:| ||  || ||| |:|:| | ::|| |||| | |  :|| || | ::||: :|||
Db    710 AYSILSEVQQALSQGSSD--SHILDLSNRFYTLIPHDFGMKKPPLLNNANSVQAKVEMLD 767

Qy    295 VLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLCQLQLLDSGAPEYKVIQTYLEQT- 353
            | :  || :: ::    |:::  ::|  :::|  ||:|::|::
Db    768 NLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVKDSEEAEIIRKYVKNTH 823

Qy    354 GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH 411
            |::|  | ||  ::|| :| |: :| :|||||||||||||:||::||||:|:|||||
Db    824 ATTHNAYDLEVVDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPP 883

Qy    412 ----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDNP 464
            :|    |||||||||::::||::||||| :    ||:|  :| | |
Db    884 EAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEAALGNMYELKHAR- 937

Qy    465 SLKSPPPGFDSVIARGHTEPDPTIQDTELELDGQQVVPVPQGQPVPCPEFSSSTFSQSEYLI 524
            | |||  | | ||: : |: ::::|     ||:: :| ||  |: |: |: :|| ::||||
Db    938 HISKLPKGHSVKGLGKTPDPS--ASTTVDG--VEVPLGTGI-SSGVNDTCLLYNEYIV 992

Qy    525 YQESQCRLRYLLEV 538
            ||||||||||||||
Db    993 YDIAQVHLKYLLKL 1006


RESULT 6
PPOL_SARPE
ID   PPOL_SARPE     STANDARD;      PRT;    996 AA.
AC   Q11208;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
OS   Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7386;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=94170813; PubMed=8125121;
RA   Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,
```

RA   de Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;
RT   "Cloning and functional expression of poly(ADP-ribose) polymerase
RT   cDNA from Sarcophaga peregrina.";
RL   Eur. J. Biochem. 220:607-614(1994).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC   nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; D16482; BAA03943.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; ZnF-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; FALSE_NEG.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   DNA_BIND   1   369   BY SIMILARITY.
FT   DOMAIN   370   507   AUTOMODIFICATION DOMAIN.
FT   DOMAIN   382   456   BRCT.
FT   DOMAIN   508   996   NAD-BINDING.
FT   ZN_FING   19   54   BY SIMILARITY.
FT   ZN_FING   126   164   BY SIMILARITY.
FT   DOMAIN   211   214   NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN   232   235   NUCLEAR LOCALIZATION SIGNAL 2ND PART.
SQ   SEQUENCE   996 AA; 113018 MW; 690DDD36E7487298 CRC64;

Query Match   23.6%; Score 674.5; DB 1; Length 996;
Best Local Similarity   33.6%; Pred. No. 2.6e-39;
Matches   179; Conservative 104; Mismatches 195; Indels 55; Gaps 22;

Qy   32 EEDPFRSTAEALKAIPAEKRI-----IRVDPTCPLSSNPGTQVY-----EDYNCTLNQTN 81
          || :: ::||| |||| ||| |: ||
Db   490 ESKSKSKSIYTKSVPKSMTLKIKDGLAVDPDSGLED--VAHVVSRNKEKYNVVLGITD 547

Qy   82 IENNNKFYIIQLLQD--SNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFRE 138
          | |||| ||| ||| |||| :|||||| | ||| ||| | || :|| |
Db   548 IQKNKNSFYKLQLLESDMKNRFWV-FRSWGRIGTTIGGNKLDNFSNLVDAIVQFKELYLE 606

Qy   139 KTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQ 197
          | | :| | | || | || : ||| | || | :||
Db   607 KSGNHFENRENFVKVAGRMYPIDIDYAEDS------KIDLSAEHDIKSKL-PLSV---Q 655

Qy   198 KLITNIFSKEMFKNTMALMDLDVKKMPLGKLSLSKQQIARGFEALEALEEALKGPTDGGQSL 257
          : | : ||| ::|| ||||||:::|||||||:::::|||||||::::|||:: ||
Db   656 DIIKLMFDVDSMKRTMMEFDLDMEKMPLGKLSQKQIQSAYKVLTEIYELIQG---GGTNA 712

Qy   258 E--ELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTV 315
          : ||||||||||| | | :|||||||| || || ||||||| ||||||||||||
Db   713 KFIDATNRFYTLIPHNFGTQSPPLLDTTEQVEQLRQMLDSLIEIECAYSLLQTEDSKADI 772

Qy   316 EEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEG 372
          : || ||| ::| : || || ::::||||| ||| ||||| || |||:
Db   773 ----NPIDKHYEQLKTKLPLDKNSEEYILLQKYVKNTHAETHKLYDLEVVDIFKVARQG 828

Qy   373 EEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGIYFASENSK 428
          :: |: |||||||:|||: :: ::|:|||||:|||| : | ::|||:|||
Db   829 EARRYKPFKKLHNRRLLWHGSRLTNFAGILSHGLKIAPPEAPVTGYMFGKGIYFADMVSK 888

Qy   429 SAGYVIGMKCGAHH--VGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP 486
          || || ||| | ||||
Db   889 SANYC----CTSHHNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFGRGRTMPNP 943

Qy   487 TQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVH 539
          : | :| :|: | :: |||:|: ||||: || :|:| ::
Db   944 SESIIRE-DG--VEIPLGKFITNDSLKSSLL-YNEFIIYDIAQVNIQYMLRMN 992

RESULT 7
PPOL_CHICK
ID   PPOL_CHICK   STANDARD;   PRT;   1011 AA.
AC   P26446;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=91340148; PubMed=1840535;
RA   Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;
RT   "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid
RT   sequence and comparison with mammalian enzyme sequences.";
RL   Gene 102:157-164(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.
RX   MEDLINE=96353841; PubMed=8755499;
RA   Ruf A., Mennissier de Murcia J., de Murcia G.M., Schulz G.E.;
RT   "Structure of the catalytic fragment of poly(AD-ribose) polymerase
RT   from chicken.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO
RP   895.
RX   MEDLINE=98191351; PubMed=9521710;
RA   Ruf A., de Murcia G.M., Schulz G.E.;
RT   "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived
RT   from crystal structures and homology modeling.";
RL   Biochemistry 37:3893-3900(1998).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.
RX   MEDLINE=98239716; PubMed=9571033;
RA   Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;
RT   "The mechanism of the elongation and branching reaction of poly(ADP-
RT   ribose) polymerase as derived from crystal structures and
RT   mutagenesis.";
RL   J. Mol. Biol. 278:57-65(1998).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR

```
CC        PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC        ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC        CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC        TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC        EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!-  CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC        nicotinamide + [ADP-D-ribosyl](N+1)-acceptor .
CC   -!-  COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!-  SUBUNIT: HOMODIMER (Potential).
CC   -!-  SUBCELLULAR LOCATION: Nuclear.
CC   -!-  MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC        AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC        FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC        THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC        AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!-  SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!-  SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC  --------------------------------------------------------------------------
CC        This SWISS-PROT entry is copyright. It is produced through a collaboration
CC        between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC        the European Bioinformatics Institute. There are no restrictions on its
CC        use by non profit institutions as long as its content is in no way
CC        modified and this statement is not removed. Usage by and for commercial
CC        entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC        or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------------
DR   EMBL; X52690; CAA36917.1; -.
DR   PIR; JH0581; JH0581.
DR   PDB; 2PAW; 27-MAY-98.
DR   PDB; 1PAX; 5-MAY-97.
DR   PDB; 2PAX; 27-MAY-98.
DR   PDB; 3PAX; 27-MAY-98.
DR   PDB; 4PAX; 27-MAY-98.
DR   PDB; 1A26; 27-MAY-98.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; 3D-structure.
FT   DNA_BIND      1    370
FT   DOMAIN      371    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      382    458       BRCT.
FT   DOMAIN      523   1011       NAD-BINDING.
FT   ZN_FING      21     56       PARP-TYPE.
FT   ZN_FING     125    162       PARP-TYPE.
FT   DOMAIN      207    209       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     403    403       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     404    404       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     410    410       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     411    411       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     432    432       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     441    441       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     442    442       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     453    453       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     454    454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     468    468       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     481    481       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     485    485       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     509    509       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     510    510       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     517    517       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    895    895       A -> R (IN REF. 1).
SQ   SEQUENCE   1011 AA;  113520 MW;  261AED3B3139144 CRC64;

Query Match             23.6%;  Score 674.5;  DB 1;  Length 1011;
Best Local Similarity   33.4%;  Pred. No. 2.7e-39;
Matches  187;  Conservative 100;  Mismatches 204;  Indels 69;  Gaps 24;

Qy       14 PW---VQTEGPE------------KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPT 58
                                     : ||      |||
Db      477 PWGAEVKTEHQEVAVDGKCSKPANMKSAGKVKEQGPSKSEKKMKLTV--KGGAAVDPD 533

Qy       59 CPLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWGRVG 112
                  :||  :|    :  |  ||    :  ::  ||||
Db      534 SGLEDS--AHVFEKGGKIFSATLGLVDIVKGTNSYKLQLLEDDRESRYWV-FRSWGRVG 590

Qy      113 EV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQE 170
               : |               ||:|||::|||  |:| ||  :| :||
Db      591 TVIGSNKLEQMPSKEDAVEHFLNLYEEKTGNSWHSK-NFTKYPKKFYPLEIDYGQDEEAV 649

Qy      171 AVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSK 230
                :  :  :  |    |     :  | | ||   :|||:||||||
Db      650 RKLTVSAG---TKSKLAKP------IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSK 700

Qy      231 QQIARGFEALEALEEALKGPTDGGQSLE--ELSSHFYTVYIPHNFGHSQPPPINSPELLQA 288
             :|  ||               |: |         ::||:| |||:||::  | : ||
Db      701 RQIQSAYSILNEVQQAV---SDGGSESQIIDLSNRFYTLIPHDFGMKKPPLLSNLEYIQA 757

Qy      289 KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQT 348
                  :| :| |    |:   |     :     |          |: | |  | ||:  |||
Db      758 KVQMLDNLLDIEVAYSLLRGGNEDGDKD---PIDINYEKLRTDIKVVDKDSEEAKIIKQ 813

Qy      349 YLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSG 405
             :|:| | |          :|  ::  :|| :| :||||||  :||:||||
Db      814 YVKNTHAATHNAYDLKVVEIFRIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSQG 873

Qy      406 LRIMPH----SGGRVGKGIYFASENSKSAGYVIGMKC---GAHHVGYMFLGEVALGREHH 458
             | | |:       | ||||:||:       :      |    |::||:|:||||
Db      874 LRIAPPEAPVTGYMFGKGIYFADMVSKSANY------CHTSQADPIGLILLGEVALGNMYE 928

Qy      459 INTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFS 518
             : :::|| : : |  |  | ||: ||| |: |||| ||  ||||  ||| |:| :
Db      929 LKNASHITKL-PKGKHSVKGLGKTAPDPTATT--TLDG--VEVPLGNGI-STGINDTCLL 982

Qy      519 QSEYLIYQESQCRLRYLLEV 538
             ::::: :  : :||::
Db      983 YNEYIVYDVAQVNLKYLLKL 1002


RESULT  8
PPOL_CRIGR
ID   PPOL_CRIGR              STANDARD;      PRT;  1012 AA.
AC   Q9R152;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21276334; PubMed=11382339;
RA   Ganesh A., Phillips E., Thacker J., Meuth M.;
RT   "Suppression of the radiation-sensitive phenotype of hamster irs1 and
RT   irs2 strains selected for resistance to 3-aminobenzamide.";
RL   Int. J. Radiat. Biol. 77:609-616(2001).
CC   -!-  FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
```

Matches 186; Conservative 100; Mismatches 205; Indels 59; Gaps 22;

```
Qy    9 APKPK---PWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNP 65
          |||       ::: ||| |::|     || ::        |  ||     ::
Db  493 APKGKSAAPSKKSKGLYKEEG--VNKSEKRMKLTLKGGAA---------VDPDSGLEHS- 540

Qy   66 GTQVYED---YNCTLNQTNIENNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV-GQSKI 119
          | ::||    :|| || | :| | :|  ||||: :::  |   | |||| || :|  |
Db  541 -AHVLEKGGKVFSATLGLVDIVKGTNSYKLQLLEQDKESRYWIFRSWGRVGTVIGSNKL 599

Qy  120 NHFTRLEDAKKDFEKKFPREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVKVDRG 178
          :|   ::|| :|| :| |  :|  || ||:|     || | :::  | ||  ::  ||
Db  600 EQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVKKL-- 653

Qy  179 PVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFE 238
          |:        | :          :::::::|||||:::|||||||||||||||::|  :
Db  654 TVKPGTK----SKLPKAVQELVGMIFDVESMKKALVEYEIDLQKMPGKLSKRQIQAAYS 709

Qy  239 ALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLAD 298
          :|:::  |:  ::||:  ||: ||::|| ||| :|||:|: |: :::| ||::::|::|
Db  710 ILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLD 768

Qy  299 IELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNH 357
          ||:| ::     :|  ::|| :|||:|: : | ||::|| ::|| |  |: | | ||:
Db  769 IEVAYSLLRGGSDDSSKD---PIDVNYEKLKTDIKVVDRDSEEAEVIRKYVKNTHATTH 824

Qy  358 RCPTLQ--HIWKVNQEGEERFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---- 411
          :   |:  |: |::||||  ::  |:|:|| |||||:: ::|| ||: |||| ||
Db  825 NAYDLEVMDIFKIEREGESQRYKPFKQLHNRRLLWHGSRTNFAGILSQGLRIAPPEAPV 884

Qy  412 SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDNPSLKS 468
          :| | | ||||:::|||::    |     ||    | ||||||||| : :: : :|  |:
Db  885 TGYMFGKIYFADMVSKSANY----CHTSQGDPIGLILLGEVALGNMYELKHAS-HISK 938

Qy  469 PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES 528
          |   || || | |||||::  |  : |:| :|| :|  |: :  :| :  ||||: |:|
Db  939 LPKGKHSVKGLGKTTPDPSASITLE---GVEVPLGTGIP-SGVNDFCLLYNEYIVYDIA 993

Qy  529 QCRLRYLLEV 538
          |  |:|||::
Db  994 QVNLKYLLKL 1003
```

```
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE (By
CC   similarity).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC       nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AF168781; AAD45817.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0
FT   DNA_BIND      1    371       BY SIMILARITY.
FT   DOMAIN      372    386       BY SIMILARITY.
FT   DOMAIN      386    522       AUTOMODIFICATION DOMAIN (BY SIMILARITY).
FT   DOMAIN      475    523       BRCT.
FT   DOMAIN      523   1012       NAD-BINDING (BY SIMILARITY).
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       PARP-TYPE.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT                                (POTENTIAL).
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT                                (POTENTIAL).
FT   MOD_RES     406    406       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     412    412       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     443    443       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     455    455       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     470    470       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     483    483       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     487    487       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     490    490       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     511    511       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     518    518       ADP-RIBOSYL[N] (POTENTIAL).
SQ   SEQUENCE   1012 AA;  112400 MW;  8D617C4DBF0CB0F7 CRC64;
```

Query Match   23.6%;
Best Local Similarity   33.8%;

Score 673.5; DB 1; Length 1012;
Pred. No. 3.1e-39;

```
RESULT 9
PPOL_HUMAN
ID   PPOL_HUMAN     STANDARD;      PRT;   1013 AA.
AC   P09874;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT OR PPOL OR PARP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90091744; PubMed=2513174;
RA   Auer B., Nagl U., Herzog H., Schneider R., Schweiger M.;
RT   "Human nuclear NAD+ ADP-ribosyltransferase(polymerizing):
RT   organization of the gene.";
RL   DNA 8:575-580(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=88076933; PubMed=3120710;
RA   Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S.,
RA   Suzuki H., Nyunoya H., Miwa M., Sugimura T.;
RT   "Nucleotide sequence of a full-length cDNA for human fibroblast
RT   poly(ADP-ribose) polymerase.";
RL   Biochem. Biophys. Res. Commun. 148:617-622(1987).
```

RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=88058958; PubMed=2824474;
RA Kurosaki T., Ushiro H., Mitsuuchi Y., Suzuki S., Matsuda M.,
RA Matsuda Y., Katunuma N., Kangawa K., Matsuo H., Hirose T.,
RA Inayama S., Shizuta Y.;
RT "Primary structure of human poly(ADP-ribose) synthetase as deduced
RT from cDNA sequence.";
RL J. Biol. Chem. 262:15990-15997(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88068596; PubMed=2891139;
RA Cherney B.W., McBride O.W., Chen D., Alkhatib H., Bhatia K.,
RA Hensley P., Smulson M.E.;
RT "cDNA sequence, protein structure, and chromosomal location of the
RT human gene for poly(ADP-ribose) polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8370-8374(1987).
RN [5]
RP SEQUENCE OF 440-1013 FROM N.A.
RX MEDLINE=87298455; PubMed=3113420;
RA Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA Miwa M.;
RT "Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and
RT expression of its gene during HL-60 cell differentiation.";
RL Biochem. Biophys. Res. Commun. 146:403-409(1987).
RN [6]
RP ERRATUM.
RA Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA Miwa M.;
RL Biochem. Biophys. Res. Commun. 148:1549-1550(1987).
RN [7]
RP SEQUENCE OF 1-94 FROM N.A.
RX MEDLINE=91099327; PubMed=2125269;
RA Yokoyama Y., Kawamoto T., Mitsuuchi Y., Kurosaki T., Toda K.,
RA Ushiro H., Terashima M., Sumimoto H., Kuribayashi I., Yamamoto Y.,
RA Maeda T., Ikeda H., Sagara Y., Shizuta Y.;
RT "Human poly(ADP-ribose) polymerase gene. Cloning of the promoter
RT region.";
RL Eur. J. Biochem. 194:521-526(1990).
RN [8]
RP SEQUENCE OF 1-39 FROM N.A.
RX MEDLINE=90211250; PubMed=2108670;
RA Ogura T., Nyunoya H., Takahashi-Masutani M., Miwa M., Sugimura T.,
RA Esumi H.;
RT "Characterization of a putative promoter region of the human
RT poly(ADP-ribose) polymerase gene: structural similarity to that of
RT the DNA polymerase beta gene.";
RL Biochem. Biophys. Res. Commun. 167:701-710(1990).
RN [9]
RP SEQUENCE OF 1-39 FROM N.A.
RA Herzog H., Schneider R., Hirsch-Kauffmann M., Schnitzer D.,
RA Schweiger M.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DDBJ databases.
RN [10]
RP ANALYSIS OF ZINC FINGERS.
RX MEDLINE=90222155; PubMed=2109322;
RA Gradwohl G., Menissier de Murcia J., Molinete M., Simonin F.,
RA Koken M.H.M., Hoeijmakers J.H.J., de Murcia G.M.;
RT "The second zinc-finger domain of poly(ADP-ribose) polymerase
RT determines specificity for single-stranded breaks in DNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2990-2994(1990).
RN [11]
RP ANALYSIS OF ZINC FINGERS.
RX MEDLINE=91072398; PubMed=2123876;
RA Ikelima M., Noguchi S., Yamashita R., Ogura T., Sugimura T.,
RA Gill D.M., Miwa M.;
RT "The zinc fingers of human poly(ADP-ribose) polymerase are
RT differentially required for the recognition of DNA breaks and nicks
RT and the consequent enzyme activation. Other structures recognize
RT intact DNA.";
RL J. Biol. Chem. 265:21907-21913(1990).
RN [12]
RP MUTAGENESIS OF CATALYTIC DOMAIN.
RX MEDLINE=91035460; PubMed=2121735;
RA Simonin F., Menissier de Murcia J., Poch O., Muller S., Gradwohl G.,
RA Molinete M., Penning C., Keith G., de Murcia G.M.;
RT "Expression and site-directed mutagenesis of the catalytic domain of
RT human poly(ADP-ribose)polymerase in Escherichia coli. Lysine 893 is
RT critical for activity.";
RL J. Biol. Chem. 265:19249-19256(1990).
RN [13]
RP NUCLEAR LOCALIZATION SIGNAL.
RX MEDLINE=92371433; PubMed=1505517;
RA Schreiber V., Molinete M., Boeuf H., de Murcia G.M.,
RA Menissier de Murcia J.;
RT "The human poly(ADP-ribose) polymerase nuclear localization signal is
RT a bipartite element functionally separate from DNA binding and
RT catalytic activity.";
RL EMBO J. 11:3263-3269(1992).
RN [14]
RP MUTAGENESIS OF CATALYTIC DOMAIN.
RX MEDLINE=97461532; PubMed=9315851;
RA Rolli V., O'Farrell M., Menissier-de Murcia J., de Murcia G.M.;
RT "Random mutagenesis of the poly(ADP-ribose) polymerase catalytic
RT domain reveals amino acids involved in polymer branching.";
RL Biochemistry 36:12147-12154(1997).
-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
-!- CATALYTIC ACTIVITY: NAD(+) + {ADP-ribosyl](N)-acceptor =
CC   nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
-!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
-!- SUBUNIT: HOMODIMER (Potential).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
-!- SIMILARITY: BELONGS TO THE PARP FAMILY.
-!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR EMBL; X16674; CAA34663.1; -.
DR EMBL; M18112; AAA60137.1; -.
DR EMBL; J03473; AAB59447.1; -.
DR EMBL; M17081; AAA51599.1; ALT_SEQ.
DR EMBL; M32721; AAA60155.1; -.
DR EMBL; M29786; AAA51663.1; -.
DR EMBL; M29545; AAA51663.1; JOINED.
DR EMBL; M29766; AAA51663.1; JOINED.
DR EMBL; M29767; AAA51663.1; JOINED.
DR EMBL; M29768; AAA51663.1; JOINED.
DR EMBL; M29769; AAA51663.1; JOINED.
DR EMBL; M29770; AAA51663.1; JOINED.
DR EMBL; M29771; AAA51663.1; JOINED.
DR EMBL; M29772; AAA51663.1; JOINED.
DR EMBL; M29773; AAA51663.1; JOINED.
DR EMBL; M29774; AAA51663.1; JOINED.
DR EMBL; M29775; AAA51663.1; JOINED.
DR EMBL; M29776; AAA51663.1; JOINED.
DR EMBL; M29777; AAA51663.1; JOINED.
DR EMBL; M29778; AAA51663.1; JOINED.
DR EMBL; M29779; AAA51663.1; JOINED.
DR EMBL; M29780; AAA51663.1; JOINED.

```
DR   EMBL; M29781; AAA51663.1; JOINED.
DR   EMBL; M29783; AAA51663.1; JOINED.
DR   EMBL; M29784; AAA51663.1; JOINED.
DR   EMBL; M29785; AAA51663.1; JOINED.
DR   EMBL; M29544; AAA51663.1; JOINED.
DR   EMBL; M29782; AAA51663.1; JOINED.
DR   EMBL; X56140; CAA39606.1; -.
DR   EMBL; X56141; CAA39606.1; JOINED.
DR   EMBL; M60436; AAA60000.1; -.
DR.  PIR; A26901; A26901.
DR   PIR; A28498; A28498.
DR.  PIR; A29725; A29725.
DR   PIR; A35635; A35635.
DR   PIR; A33321; A33321.
DR   PIR; B33321; B33321.
DR   PIR; A39976; A39976.
DR   PIR; S14010; S14010.
DR   HSSP; P26446; 1A26.
DR   Aarhus/Ghent-2DPAGE; 1620; NEPHGE.
DR   MIM; 173870; -.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
```

```
PPOL_RAT
ID   PPOL_RAT      STANDARD;      PRT;   1013 AA.
AC   P27008; O35937;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN   ADPRT.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Monocytes;
RX   MEDLINE=98046546; PubMed=9385436;
RA   Beneke S., Meyer R., Buerkle A.;
RT   "Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly
RT   (ADP-ribose) polymerase.";
RL   Biochem. Mol. Biol. Int. 43:755-761(1997).
RN   [2]
RP   REVISION TO 811.
RA   Beneke S., Meyer R., Buerkle A.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-11 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=92290013; PubMed=1601134;
RA   Potvin F., Thibodeau J.V., Kirkland J.B., Dandenault B.,
RA   Duchaine C., Poirier G.G.;
RT   "Structural analysis of the putative regulatory region of the rat
RT   gene encoding poly(ADP-ribose) polymerase.";
RL   FEBS Lett. 302:269-273(1992).
RN   [4]
RP   SEQUENCE OF 514-1013 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;
RX   MEDLINE=90027702; PubMed=2508731;
RA   Thibodeau J., Gradwohl G., Dumas C., Clairoux-Moreau S., Brunet G.;
RT   "Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase
RT   catalytic domain and analysis of mRNA levels during the cell cycle.";
RL   Biochem. Cell Biol. 67:653-660(1989).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U94340; AAC53544.1; -.
DR   EMBL; X65496; CAA46477.1; -.
DR   EMBL; X65497; CAA46478.1; ALT_INIT.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
```

```
Query Match             23.5%;   Score 670;  DB 1;  Length 1013;
Best Local Similarity   33.6%;   Pred. No. 5.5e-39;
Matches  186;  Conservative  101;  Mismatches  195;  Indels  72;  Gaps  23;

Qy      7 AMAPKPKPWVQTEGPEKKKGR-----QAGREEDP---FRSTAEALKAIPAEKRIIRVDPT 58
            ::| | |::  ||| |  :| :|              | |::  |:|   ||:||| |
Db    501 ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLE---------------- 546

Qy     59 CPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWRGVGEV-G 115
            | ::|  :|| :| || |||||:|||:|||||||   ||:|:||||||||||||:| |
Db    547 ------KGGKV---FSATLGLVDIVKGTNSYYKLQLLEDDKENRYW-IFRSWGRVGTVIG 596

Qy    116 QSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVK 174
            |:|      :||| :|| | |:|| | || |   ::| : :|  |||  |||||||:|||
Db    597 SNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVKK 653

Qy    175 VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA 234
            | | |  | | ::|   |||||||:||:|:||||||||||||||||||||||||||||||
Db    654 LTVNP-GTKSKLFKP-----VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQ 706

Qy    235 RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL 294
            | |:|:  ||::| | |::: |::  ||| ||| |  |  ::   ::||| | :|:| :|
Db    707 AAYSILSEVQQAVSQGSGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLD 765

Qy    295 VLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT- 353
            || || ||:|  | :| :|  |   | |: |:  ::|  ||: | ::|||: |: |:: |
Db    766 NLLDIEVAYSLLRGGSDSSKD---PIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTH 821

Qy    354 GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH 411
            ||:| |||:|  :: : || :  |||||| | ||| |:||::| || |||:||||:|||:
Db    822 ATTHNAYDLEVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPP 881

Qy    412 ---SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDNP 464
               |::||||||| |||| :| : :|||||:|||:       | |||||| |:  ||:  :
Db    882 EAPVVGYMFGKGIYFADMVSKSANY------CHTSQGDPIGLILLGEVALGNMYELKHAS- 935

Qy    465 SLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYLI 524
             :| |  |: |::  || |: | ::||:|: || | ::|  |  ::  ::||:::||:
Db    936 HISKLPKGKHSVKGLGKTTPDPS--ANISLDG--VDVPLGTGI-SSGVNDTSLLYNEYIV 990

Qy    525 YQESQCRLRYLLEV 538
            |: | || |:| ||
Db    991 YDIAQVNLKYLLKL 1004
```

RESULT 10

```
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc.
FT   INIT_MET      0      0
FT   DNA_BIND      1    372       BY SIMILARITY.
FT   DOMAIN      385    461       BRCT.
FT   DOMAIN      373    523       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      524   1013       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       PARP-TYPE.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES       1      1       BLOCKED (BY SIMILARITY).
FT   MOD_RES     407    407       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     413    413       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     435    435       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     437    437       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445    445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     456    456       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     484    484       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     488    488       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     491    491       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     513    513       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     519    519       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    638    638       Y -> H (IN REF. 4).
FT   CONFLICT    641    641       E -> A (IN REF. 4).
FT   CONFLICT    752    752       N -> D (IN REF. 4).
SQ   SEQUENCE   1013 AA;  112529 MW;  AA566F2B29BE97C0 CRC64;

Query Match             23.3%;
Best Local Similarity   33.5%;           Score 665.5;  DB 1;  Length 1013;
Matches  186;  Conservative  99;  Mismatches  203;  Indels  67;  Gaps  23;
            Pred. No. 1.1e-38;

Qy    8  MAPKPKPWVQPEGPEKKGKGRQAGREEDPFRSTAEEALKAIPAEKRI------IRVDPTCP   60
Db  493  VVPKGK-----SAAPSKKSKGAVKEEGVNKS--------EKRMKLTLKGGAAVDPDSG    537

Qy   61  LSSNPGTQVYED----YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEV-   114
Db  538  LEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKQLLESDKESRYWIFRSWGRVGTVI   595

Qy  115  GQSKINHFTRLEDAKKDFEKKFREKTKNNAAERDHFVSHPGKYTLIEVQ-AEDEAQEAVV   173
Db  596  GSNKLEQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVK   652

Qy  174  KVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQI   233
Db  653  KL---AVKPGTK-----SKLPKPVQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQI   705

Qy  234  ARGFEALEALEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDML   293
Db  706  QAAYSILSEVQQAVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNTDSVQAKVEML   764

Qy  294  LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT   353
Db  765  DNLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEVIRKVKNT   820

Qy  354  -GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP   410

Db  821  HATTHNAYDLEVIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQCLRIAP   880
Qy  411  H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDN   463
Db  881  PEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEVALGNMVELKHAS   935

Qy  464  PSLKSPPPGFDSVIARGHTEPDTQDTELELDGQQVVVPQGQVPCPEFSSSTFSQSEYL   523
Db  936  -HISKLPKGKHSVKGLGKTAPDFS--ASITLDG--VEVPLGTGIP-SGVNDTCLLYNEYI   989

Qy  524  IYQESQCRLRYLLEV   538
Db  990  VYDIAQVNLKYLLKL   1004
```

```
RESULT 11
PPOL_XENLA
ID   PPOL_XENLA    STANDARD;    PRT;    998 AA.
AC   P31669;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Saulier-Le Drean B.M.;
RL   Thesis (1992), University of Rennes, France.
RN   [2]
RP   SEQUENCE OF 742-876 FROM N.A.
RX   MEDLINE=9277538; PubMed=8503897;
RA   Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada M.,
RA   Miwa M.;
RT   "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)
RT   polymerase from Xenopus laevis and cherry salmon using heterologous
RT   oligonucleotide consensus sequences.";
RL   Biochem. Biophys. Res. Commun. 193:119-125(1993).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,
CC       AND BRAIN. LOW IN LIVER.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
```

```
DR   EMBL; Z12139; CAA78126.1; -.
DR   EMBL; D13810; BAA02966.1; -.
```

```
DR   PIR; S31735; S31735.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; zinc-finger; zinc.
FT   NON_TER       1       1
FT   DNA_BIND     <1     356
FT   DOMAIN      357     507       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      369     445       BRCT.
FT   DOMAIN      508     998       NAD-BINDING.
FT   ZN_FING       8      43       PARP-TYPE.
FT   ZN_FING     111     148       PARP-TYPE.
FT   DOMAIN      193     195       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      207     212       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     391     391       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     397     397       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     419     419       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     428     428       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     429     429       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     445     445       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     447     447       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     454     454       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     467     467       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     471     471       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     477     477       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     495     495       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     496     496       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     503     503       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    746     746       Q -> E (IN REF. 2).
SQ   SEQUENCE    998 AA;   111126 MW;  F5A25E4A3366BAE7 CRC64;
```

```
Query Match             22.9%;   Score 653;  DB 1;  Length 998;
Best Local Similarity   32.5%;   Pred. No. 8.2e-38;
Matches  177;  Conservative 102;  Mismatches 214;  Indels 52;  Gaps 20;

Qy     10  PKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQV   69
               | | :: :|::  |:  :||:|    | ::  :|: | ::: :|
Db    480  PSSGPVAGKSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDSEL---EDSCHVLETGG---  533

Qy     70  YEDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEV-GQSKINIFTRLED  127
               | :: ||  :|:|   : : |:|||   | ||:||:| |  | |||| ::|
Db    534  -KIFSATLGLVDITRGTNSYYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKED  592

Qy    128  AKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV  187
           | ::  ||||::||   |  :|| |  |: :| |: : ::|  :  | |  |  |   ||
Db    593  AIEHFLNLYQDKTGNAW-HSPNFTKYPKKFYPLEIDYGGE-EDVVKKLSVG-AGTKSKLA  649

Qy    188  QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL  247
               :|  || | |   | ::| ::|:|||||||  | |  ||  |:| : :::::|
Db    650  KP-----VQELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAV  703

Qy    248  KGPTDGGQSLEE-----LSSHFYIVIPHNFGHSQOPPPINSPELLQAKKDMLLVLADIELA  302
           || | || |::       |:  |: ::|:::| | |:|| | | ||:| || :|:| |:|
Db    704  -------SESLSEARLLDLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVA  757

Qy    303  QAL-QAVSEQEKTVEEVPHPLDRDYQLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRC  359
           || | | :    | ||| | |  :| ::: : |: || :   ::   |||    |: |
Db    758  YSLLRGGADDGEK------DPIDVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNA  811

Qy    360  PTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SG  413
```

RESULT 12

```
PPOL_MOUSE
ID   PPOL_MOUSE     STANDARD;       PRT;   1012 AA.
AC   P11103; Q93LX4; Q9QVQ3;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)
DE   ADP-ribosyltransferase-1) (Poly[ADP-ribose) synthetase-1) (mSPARP).
GN   ADPRT OR ADPRT1 OR ADPRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=BXSB;
RX   MEDLINE=89263780; PubMed=2498841;
RA   Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;
RT   "Sequence and organization of the mouse poly (ADP-ribose) polymerase
RT   gene.";
RL   Nucleic Acids Res. 17:3387-3401(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=129/Sv X C57BL/6; TISSUE=Fibroblast;
RX   MEDLINE=20270268; PubMed=10809783;
RA   Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;
RT   "Characterization of sPARP-1. An alternative product of PARP-1 gene
RT   with poly(ADP-ribose) polymerase activity independent of DNA strand
RT   breaks.";
RL   J. Biol. Chem. 275:15504-15511(2000).
RN   [3]
RP   KNOCK-OUT.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jaeger S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC       PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC       ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC       CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC       TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC       EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC       nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (Potential).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC       short form/sPARP-1; may be produced by alternative initiation.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC       AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC       FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC       THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
```

```
Db    812  YDLEVLEIFKIDREGEYQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTG  871
               :| ::::|||  |:::||||  :  :||::|||||
Qy    414  GRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGF  473
               |||||  |  :|||||||  : | ::||:|||
Db    872  YMFGKGIYFADMVSKSANYCHAMP--GSPIGLILLGEVALGNMHELKAASQITKL-PKGK  928
               || ||| :|| |::|||:|| |    ::||||
Qy    474  DSVIARGHTEPDTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLR  533
               :: || || :|:: |  ::|  | :| : || :|::||||
Db    929  HSVKGLGRTAPDPS--ATVQLDG--VDVPLGKGTSA-NISDFSLLYNEYIVIVDIAQVNLK  983

Qy    534  YLLEV  538
           |||:|
Db    984  YLLKL  988
```

```
Query Match                22.9%;  Score 653;  DB 1;  Length 1012;
Best Local Similarity      32.3%;  Pred. No. 8.4e-38;
Matches 186; Conservative 102; Mismatches 211; Indels 76; Gaps 23;

Qy    2 SLLFLAMAPKPKW---VQTEGPE-----------KKKGRQAGREEDPFRSTAEALKAIP 47
          :   | |      :   | :           | ::|  |   :|: |  :|
Db  467 SLQDLLSAHSLSPWGAEVKAEPGEVVAPRGKSAAPSKKSKGCFKEEGVNKS-------- 517

Qy   48 AEKRI------IRVDPTCPLSSNPGTQVYED----YNCTLNQTNIENNNKFYIIQLLQ 96
          ||           |:|         |       ||   :|: | :||::::|::
Db  518 -EKRMKLTLKGGAAVDPDSGLEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLE 574

Qy   97 DSNRF-FTCWNRWGRVGEV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHP 154
         | ::  :   |:|:| ||   | :  | :|  | |  :   |: |::|: |::  :| :
Db  575 DDKESRYWIFRSWGRLGTVIGSNKLEQMPSKEAVEQFMKLYEKTGNAWHSK-NFTKYP 633

Qy  155 GKYTLIEVQ-AEDEAQEAVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTM 213
         |::  |: :   :: |:  |||    :|  :|| :      |: :  :||:|::|
Db  634 KKFYPLEIDYGQDE--EAVKKLTVKP-GTKSKLPKP-------VQELVGMIFDVDSMKKAL 684

Qy  214 ALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFG 273
         ::   : :|||||||| :||   :| | :     ::     || :|  |:|| ||||:||
Db  685 VEYEIDLQKMPLGKLSRRQIQAAYSILSEVQQPVSQGSSESQIL-DLSNRFYTLIPHDFG 743

Qy  274 HSQPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQL 333
            :    ||  ::||::|| | ||:|:|:  :|:||: :|     |:|| | :  |::
Db  744 MKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDI 799

Qy  334 QLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ-HIWKVNQEGEEDRFQAHSKLGNRKLLW 390
         ::|  ::||  : ||: :|   |::      |:|  :|::|||||:|  | |  :|:|||
Db  800 KVVDRDSEEAEVIRKYVKNTHATTHNAYDLEVIDIFKIEREGESQRYKPFRQLHNRLLW 859

Qy  391 HGTNMAVVAAILTSGLRIMPH---SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---V 443
         ||:   :  :  |:||||:|      | ::||||||:  ||::|      ||           |
Db  860 HGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPI 914

Qy  444 GYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQ 503
         | : |||||| :  :         |  | | |   |||: ||:: ::|  |  ::|:||
Db  915 GLIMLGEVALGNMYELKHAS-HISKLPKGKHSVKGLGKTTPDPSASITLE---GVEVPL 969

Qy  504 GQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEV 538
         | :   |  | :| :||| ::|: ::  | :||:
Db  970 GTGIP-SGVNDTALLYNEYIVYDIAQVNLKYLLKL 1003
```

```
RESULT 13
PPOL_DROME
ID   PPOL_DROME     STANDARD;     PRT;     994 AA.
AC   P35875; Q9W5Q5; Q9W5S1;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+) ADP-
DE   ribosyltransferase) (Poly[ADP-ribose] synthetase).
GN   PARP OR CG17696/CG17718.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93234521; PubMed=8475096;
RA   Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,
RA   Sugimura T., Miwa M.;
RT   "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:
RT   leucine zipper in the auto-modification domain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).
RN   [2]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND
RP   TISSUE SPECIFICITY.
RC   STRAIN=CANTON-S;
RX   MEDLINE=98234380; PubMed=9565614;
```

```
CC       AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   ------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------
DR   EMBL; X14206; CAA32421.1; -.
DR   EMBL; AF126717; AAF61293.1; ALT_INIT.
DR   PIR; S04200; S04200.
DR   HSSP; P26446; 1A26.
DR   MGD; MGI:1340806; Adprt1.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; ZnF-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative initiation.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1   1012       POLY [ADP-RIBOSE] POLYMERASE-1, LONG
FT                                ISOFORM.
FT   CHAIN       521   1012       POLY [ADP-RIBOSE] POLYMERASE-1, SHORT
FT                                ISOFORM.
FT   INIT_MET    521    521       FOR SHORT ISOFORM.
FT   DNA_BIND    521    371
FT   DOMAIN      372    522       AUTOMODIFICATION DOMAIN.
FT   DOMAIN      384    460       BRCT.
FT   DOMAIN      523   1012       NAD-BINDING.
FT   ZN_FING      20     55       PARP-TYPE.
FT   ZN_FING     124    161       PARP-TYPE.
FT   DOMAIN      206    208       NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN      220    225       NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   MOD_RES     406    406       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     412    412       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     434    434       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     436    436       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     443    443       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     444    444       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     447    447       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     455    455       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     483    483       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     487    487       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     490    490       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     511    511       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     512    512       ADP-RIBOSYL[N] (POTENTIAL).
FT   MOD_RES     518    518       ADP-RIBOSYL[N] (POTENTIAL).
FT   CONFLICT    590    590       L -> V (IN REF. 2).
FT   CONFLICT    607    607       E -> D (IN REF. 2).
FT   CONFLICT    611    611       Q -> H (IN REF. 2).
FT   CONFLICT    628    628       N -> D (IN REF. 2).
FT   CONFLICT    678    678       D -> E (IN REF. 2).
FT   CONFLICT    702    702       R -> E (IN REF. 3).
FT   CONFLICT    716    716       Q -> E (IN REF. 2).
FT   CONFLICT    757    757       Q -> L (IN REF. 3).
FT   CONFLICT    856    856       R -> F (IN REF. 3).
FT   CONFLICT    981    981       A -> C (IN REF. 2).
SQ   SEQUENCE   1012 AA;  112968 MW;  4354C3E5F01B9439 CRC64;
```

RA   Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;
RT   "Genomic organization of Drosophila poly(ADP-ribose) polymerase and
RT   distribution of its mRNA during development.";
RL   J. Biol. Chem. 273:11881-11886(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (LONG ISOFORM).
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC   PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC   ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC   CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC   TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC   EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC   SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES
CC   OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM
CC   IN LATER EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN
CC   EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.
CC   -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC   AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC   FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC   THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC   AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; D13806; BAA02964.1; -.
DR   EMBL; AF051548; AAC24518.1; -.
DR   EMBL; AF051544; AAC24518.1; JOINED.
DR   EMBL; AF051545; AAC24518.1; JOINED.
DR   EMBL; AF051546; AAC24518.1; JOINED.
DR   EMBL; AF051547; AAC24518.1; JOINED.
DR   EMBL; AE002935; AAF45400.1; -.
DR   EMBL; AE002666; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE002892; AAF45445.2; ALT_SEQ.
DR   PIR; A47474; A47474.
DR   HSSP; P26446; 1A26.
DR   FlyBase; FBgn0010247; Parp.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00347; PARP_ZN_FINGER_1; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW   Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW   ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.
FT   DNA_BIND   1   367
FT   DOMAIN   368   507   AUTOMODIFICATION DOMAIN.
FT   DOMAIN   380   454   BRCT.
FT   DOMAIN   508   994   NAD-BINDING.
FT   ZN_FING   19   54   PARP-TYPE.
FT   ZN_FING   123   161   PARP-TYPE.
FT   DOMAIN   208   210   NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT   DOMAIN   223   228   NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT   VARSPLIC   376   564   MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   994 AA; 113791 MW; ACA85A270DD29E08 CRC64;

Query Match   22.4%; Score 638; DB 1; Length 994;
Best Local Similarity   32.2%; Pred. No. 9.1e-37;
Matches 168; Conservative 100; Mismatches 195; Indels 58; Gaps 19;

Qy   44 KAIPAEKRI-----IRVDPTCPLSSNPGTQVYED---YNCTLNQTIENNNKFVIIQL 94
          |:|:| ||| ||::||:: | ::: :|| | :||
Db   500 KSMPVSRTFKVKDGLAVDPDSGLEDI-AHYVDSNNKYSVVLGLTDIQRNKNSYKVQL 557

Qy   95 LQ-DSNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFREKTKNWAERDHFVS 152
          |:| ||:||:   :||::| ::|||| | |:|::   :|:|:|
Db   558 LKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFKEIYADKTGNEYEQRDNFVK 617

Qy   153 HPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNT 212
          |: ||:::| : :|| ::|| | |
Db   618 RTGRMYPIEIQYDDD--QKLVKHE-------SHFTSKLEISVQNLIKLIFDIDSMNKT 667

Qy   213 MALMDLDVKKMPLGKLSKQQIARGFEALEAEEALKGPTDGGQSLEELSSHFYTVIPHNF 272
          : ||||||||||||| : | : |:: || ||:|:| | :|| || |
Db   668 LMEFHIDMDKMPLGKLSAHQIQSAYRVKEIYNVLECGSNTAK-LIDATNRFYTLIPHNF 726

Qy   273 GHSQPPPINSPELLQAKKDMLLVLADIELAQOLAQVSEQEXTVEEVPHPLDRDYQLLKCQ 332
          | | | || ||: | :||: :| :| : || | :|| ||| ::::
Db   727 GVQLPLIETHQQIEDLRQMLDSLAEIEVAYSI----IKSEDVSDACNPLDNHYAQIKTQ 782

Qy   333 LQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLL 389
          | | |||:|:|:|:|| |:| |:||::| | :|| ::|| :|||:|:|:|:|::||
Db   783 LVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARRFKPFFKLHNRKLL 842

```
Qy    390  WHGTNMAVVAAILTSGLRI---MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY 445
                 ||||  |||  | :|| ||||||| |||| |  ||  :|
Db    843  WHGSRLTNFVGILSHGLRIAPPEAPTGVMFGKGIYFADMVSKSANYCCTSQ--QNSTGL 900

Qy    446  MFLGEVALG------REHHIN--TDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQ 497
                  |  :|      | :||   :||  |||| |||  ||| |||  |||
Db    901  MLLSEVALGDMMECTSAKYINKLSNNK------HSCFGRGRTMPDPTK-SYIRSDG- 949

Qy    498  QVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEV 538
            - |:|:: ||   :: | |   |  :||||||||||| :||||
Db    950  -VEIPYGETI-TDEHLKSSLLYNEYIVYDVAQVNIQYLFRM 988
```

```
RESULT  14
YQN4_CAEEL
ID    YQN4_CAEEL      STANDARD;        PRT;       538 AA.
AC    Q09525;
DT    01-FEB-1996 (Rel. 33, Created)
DT    01-FEB-1996 (Rel. 33, Last sequence update)
DT    01-FEB-1996 (Rel. 33, Last annotation update)
DE    Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.
GN    E02H1.4.
OS    Caenorhabditis elegans.
OC    Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC    Rhabditidae; Peloderinae; Caenorhabditis.
OX    NCBI_TaxID=6239;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=BRISTOL N2;
RA    Smith A.;
RL    Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC    -!- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-
CC        RIBOSYLTRANSFERASE (EC 2.4.2.30).
CC    --------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute.  There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------------
DR    EMBL; Z47075; CAA87379.1; -.
DR    HSSP; P26446; 1A26.
DR    WormPep; E02H1.4; CE01539.
DR    InterPro; IPR001290; PARP.
DR    InterPro; IPR004102; PARP_reg.
DR    Pfam; PF00644; PARP; 1.
DR    Pfam; PF02877; PARP_reg; 1.
KW    Hypothetical protein.
SQ    SEQUENCE  538 AA;  61268 MW;  3144E25465FC7341 CRC64;
```

```
Query Match             13.8%;  Score 393.5;  DB 1;  Length 538;
Best Local Similarity   26.1%;  Pred. No. 4.3e-20;
Matches  149;  Conservative  90;  Mismatches 172;  Indels 159;  Gaps  27;

Qy     73  YNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDF 132
               || : || |||  ||||||  ::|| |::|   |::| | | |  |||:|
Db     12  YKVHLCKTNIAQNNNKFYDMELLDEGGDFIVKLIN-GRIGYRGVTQLKDFDDLDRAKKFF 70

Qy    133  EKKFREKTKNNWAERDHFVSHPGKYTLIEV--------QAEDEAQEAVVKVD-------- 176
            |||| |  ||  |||:  |  :| |  |:||        ||||||| ||||
Db     71  ESKFYEKTHLHWEERDD-EPVPNKYAVVELATNARQTEKEVKKEEPEPEPKVDEKNTRGR 129

Qy    177  --RGPV--RTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDV----KKM---- 223
              ||:    |:|| | |: ||   :| || | ||   |::|  || |||
Db    130  KKRGIVKEKKEIKKEEEP--VEEVNEKL------KELMK---CICDEDVHGLLKQLKFN 178

Qy    224  -----PLGKLSKQQIARGFEALEALEEALKGPT---------------DGGGSLEE 259
                 || |:  || |:|:|  | | |: ||||               ||
Db    179  EAFGRPIDCLSLAQLTTGYEILSKIEESIGGKSAARRSTRGPRVADRVLAVKSDGPSLHD 238
```

```
Qy    260  LSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVL-----ADIELAQALQAVSEQEKT 314
            :: :: |||| ||| |||||||| || ||:||| :||     : || :|| || ||
Db    239  INK-YYSLIPHSFGFCVPPKIDSHAKIQAERELLDALKGSIEASLELKDLKKTASSKD- 295

Qy    315  VEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQ---HIWKVN- 369
            |  | | | |  |||  |  || ||| ||| | :|| :| :|  ||:|      :|:||
Db    296  -----IYQRLYERLPCHLEPVSE------------ETAGKIGDCLAMRGPTHCYKLSLI 337

Qy    370  -----------------QEGEEDRFQAHSKLG----NRKLLWHGTNMAVVAAIL 402
                             |  |||::  :| ||    ::||| |||||||| |||
Db    338  DAFELKDPNEIPTEAPVEVQEVPKKRGKSTKTAAPTVPPPTKRLLWHGTRVTNVFSIL 397

Qy    403  TSGLR--IMPHSGGRVGKGIYFASENSKSAGYVIGMKC-GAHHVGYMFLGEVALGR---- 455
             | ||   ||| |||| ||||||||||||||| | | |   |:| |: |:|  :|
Db    398  MNGLQFPVGDRCGLMFGNGVYFANVPFKSANYC---CPEASKRVFMLLCEVETANPLVL 453

Qy    456  -EHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQVVVPQGQPVVPCPEFSS 514
             : |:|  :| |||||||||||||||||||||| |||  ||| :|| |:|  |:||| |
Db    454  YESEIDADEKMEKAKK--TSVTAAGKHTP---RDT-VEING-------IPAFKSNL 496

Qy    515  STFSQ------SEYLIYQESQCRLRYLLEV 538
            | | |      |||||||:|:|||||||||
Db    497  ETIEEETRLLYDEYVMFNKEHFKIKYVVEV 526
```

```
RESULT  15
PPOV_HUMAN
ID    PPOV_HUMAN      STANDARD;        PRT;      1724 AA.
AC    Q9UKK3; O75903; Q9H1M6; Q14682;
DT    16-OCT-2001 (Rel. 40, Created)
DT    16-OCT-2001 (Rel. 40, Last sequence update)
DT    01-MAR-2002 (Rel. 41, Last annotation update)
DE    Vault poly(ADP-ribose) polymerase (EC 2.4.2.30) (VPARP) (193-kDa vault
DE    protein) (PARP-related/Talphai-related H5/proline-rich) (PH5P).
GN    ADPRTL1 OR PARPL OR KIAA0177.
OS    Homo sapiens (Human).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX    NCBI_TaxID=9606;
RN    [1]
RP    SEQUENCE FROM N.A., AND SEQUENCE OF 306-319.
RX    MEDLINE=99408776; PubMed=10477748;
RA    Kickhoefer V.A., Siva A.C., Kedersha N.L., Inman E.M., Ruland C.,
RA    Streuli M., Rome L.H.;
RT    "The 193 kDa vault protein, VPARP, is a novel poly(ADP-ribose)
RT    polymerase.";
RL    J. Cell Biol. 146:917-928(1999).
RN    [2]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Thymus;
RX    MEDLINE=20112770; PubMed=10644454;
RA    Still I.H., Vince P., Cowell J.K.;
RT    "Identification of a novel gene (ADPRTL1) encoding a potential
RT    poly(ADP-ribosyl)transferase protein.";
RL    Genomics 62:533-536(1999).
RN    [3]
RP    SEQUENCE FROM N.A.
RA    Tromans A.;
RL    Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN    [4]
RP    SEQUENCE OF 94-1724 FROM N.A.
RC    TISSUE=Bone marrow;
RX    MEDLINE=96281124; PubMed=8724849;
RA    Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT    "Prediction of the coding sequences of unidentified human genes. V.
RT    The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT    analysis of cDNA clones from human cell line KG-1.";
RL    DNA Res. 3:17-24(1996).
RN    [5]
RP    DISCUSSION OF SEQUENCE.
RX    MEDLINE=99198702; PubMed=10100603;
RA    Jean L., Risler J.-L., Nagase T., Coulouarn C., Nomura N.,
RA    Salier J.-P.;
```

RT   "The nuclear protein PH5P of the inter-alpha-inhibitor superfamily: a
RT   missing link between poly(ADP-ribose)polymerase and the
RT   inter-alpha-inhibitor family and a novel actor of DNA repair?";
RL   FEBS Lett. 446:6-8(1999).
CC   -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =
CC   nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.
CC   -!- SUBUNIT: BINDS TO THE MAJOR VAULT PROTEIN (MPV). THE VAULT, A
CC   LARGE RIBONUCLEOPROTEIN COMPLEX, CONTAINS THE 100-kDA MPV AND 2
CC   PROTEINS OF 193 AND 240-kDA.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, BUT IS ALSO FOUND IN THE
CC   NUCLEUS, ASSOCIATED WITH MITOTIC SPINDLES.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; THE HIGHEST LEVELS ARE IN
CC   THE KIDNEY; ALSO DETECTED IN HEART, PLACENTA, LUNG, LIVER,
CC   SKELETAL MUSCLE, SPLEEN, LEUKOCYTES AND PANCREAS.
CC   -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AF158255; AAD47250.1; -.
DR   EMBL; AF057160; AAC62491.1; -.
DR   EMBL; AL359763; CAC21562.1; -.
DR   EMBL; D79999; BAA11494.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR002035; VWFA.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00292; BRCT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW   ADP-ribosylation; Ribonucleoprotein.
FT   DOMAIN   1   94   BRCT.
FT   DOMAIN   256   569   NAD-BINDING (BY SIMILARITY).
FT   DOMAIN   876   1046   VWFA.
FT   DOMAIN   1562   1724   INTERACTION WITH THE MAJOR VAULT PROTEIN.
FT   DOMAIN   19   25   NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT   (POTENTIAL).
FT   DOMAIN   1237   1249   NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT   (POTENTIAL).
FT   CONFLICT   519   519   S -> P (IN REF. 1).
FT   CONFLICT   873   873   S -> N (IN REF. 2 AND 4).
FT   CONFLICT   897   897   Q -> E (IN REF. 2 AND 4).
FT   CONFLICT   899   899   A -> T (IN REF. 1).
FT   CONFLICT   936   936   A -> T (IN REF. 1).
FT   CONFLICT   936   936   A -> M (IN REF. 3).
FT   CONFLICT   1012   1012   V -> I (IN REF. 4).
FT   CONFLICT   1065   1065   A -> V (IN REF. 3 AND 4).
FT   CONFLICT   1080   1080   R -> L (IN REF. 3).
FT   CONFLICT   1108   1108   C -> R (IN REF. 3).
FT   CONFLICT   1265   1265   A -> A (IN REF. 2 AND 4).
FT   CONFLICT   1280   1280   G -> R (IN REF. 2 AND 4).
FT   CONFLICT   1328   1328   T -> P (IN REF. 3).
FT   CONFLICT   1331   1331   T -> A (IN REF. 3).
FT   CONFLICT   1394   1394   A -> S (IN REF. 3).
FT   CONFLICT   1459   1459   Y -> L (IN REF. 3).
FT   CONFLICT   1550   1550   P -> L (IN REF. 3).
FT   CONFLICT   1555   1555   V -> L (IN REF. 1).
FT   CONFLICT   1564   1564   T -> I (IN REF. 3).
FT   CONFLICT   1656   1656   P -> A (IN REF. 3).
SQ   SEQUENCE   1724 AA; 192587 MW; A009F3493460EDC CRC64;

```
Query Match             9.0%;  Score 258;  DB 1;  Length 1724;
Best Local Similarity   22.8%;  Pred. No. 5.9e-10;
Matches  127;  Conservative  92;  Mismatches  193;  Indels  146;  Gaps  25;

Qy      49  EKRIIRV---DPTCPLSSNP------GTQVYEDYNCTLNQTNIEN------NNNKFYI  91
            ||||:|| | || |    :: | :    |:|  :  |  : :|:|
Db      83  EKRLLDVKNYDPYKPLDITPPPDQKASSSEVKTEGLCPDSATEEEDTVELTEFGMQNVEI 142

Qy      92  IQLLQDSNRFTCWNRWGRVG-EVGQSKI-----------------NHF-------TR-  124
            | ||    :|  ||  ||  :
Db     143  PHLPQDFE--VAKYNTLEKVGMEGGQEAVVVELQCSRDSRDCPFLISSHFLLDDGMETRR 200

Qy     125  ------LEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDR  177
                 | :        :  |      : :||
Db     201  QFAIKKTSEDASEYFENYIEELKKQGFLLREHF------TPEATQLASEQLQALL---  249

Qy     178  GPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLD-VKKMPLGKLSKQQIARG 236
            |  :  :     |        :  :  |  :  :  :     |:|  ::
Db     250  ----LEEVMNSSTLSQEVSDLVEMIWAE------ALGHLEHMLLKPVNRISLNDVSKA  297

Qy     237  FEALEALEEALK-GPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLV  295
            | :  ||| |:|| || | ||  :| ||  |:||    ||  :  :  | |   |:|
Db     298  EGILLLVKAALKNGET--AEQLQKMMTEFYRLIPHK--GTMPKEVNL-GLLAKKADLCQL 352

Qy     296  LADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEY-KVIQTYLEQTG 354
              | :                :       |:|:::||  ||:   |:
Db     353  IRDM------VNVCETNLSKPNPPSLAKYRALRCKIEHVEQNTEFLRVRKEVLQNHH   404

Qy     355  SNHRCPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGL--RIMPH  411
            || | |::|:|  :  |||||||| :  |:|     :
Db     405  SKSPVDVLQ-IFRVGRVNETTEFL-SKLGNVRPLLHGSPVQNIVGILCRGLLLPKVVED  461

Qy     412  SG------GRVGKGIYFASENSKSAGYVIGMKCGAHHVG------YMFLGEVALGREHHI 459
            ||            | : :| | |  |   | || |         :| ||| | |||:
Db     462  RGVQRTDVGNLGSGIYFSDSLSTSIKY------SHPGETDGTRLLLICDVALGKCMDL   513

Qy     460  NTDNPSLKSPPPGFDSVIARGHTEPDTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQ  519
            : : | :  || |       |   :|  |::
Db     514  HEKDFSLTEAPPGYDSVHGVSQTASVTTD---------------------FED      545

Qy     520  SEYLIYQESQCRLRYLLE  537
            |:::|:|   |  ::|   :
Db     546  DEFVVYKTNQVKMKYIIK  563
```

Search completed: August 29, 2002, 08:02:19
Job time: 296 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

```
              GenCore version 4.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 29, 2002, 08:01:31 ; Search time 117.38 Seconds
                                              (without alignments)
                                              795.853 Million cell updates/sec

Title:           US-09-701-586B-6
Perfect score:   2854
Sequence:        1 MSLLFLAMAPKPKPWVQTEG.........EYLIYQESQCRLRYLLEVHL 540

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters:     562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_19:*
             1:  sp_archea:*
             2:  sp_bacteria:*
             3:  sp_fungi:*
             4:  sp_human:*
             5:  sp_invertebrate:*
             6:  sp_mammal:*
             7:  sp_mhc:*
             8:  sp_organelle:*
             9:  sp_phage:*
            10:  sp_plant:*
            11:  sp_rodent:*
            12:  sp_virus:*
            13:  sp_vertebrate:*
            14:  sp_unclassified:*
            15:  sp_rvirus:*
            16:  sp_bacteriap:*
            17:  sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2823	98.9	533	4	Q96CG2	Q96cg2 homo sapien
2	2249.5	78.8	528	11	Q91YR6	Q91yr6 mus musculu
3	828	29.0	612	5	Q9TX06	Q9tx06 dictyosteli
4	694	24.3	635	10	081294	081294 arabidopsis
5	670.5	23.5	653	10	050017	050017 zea mays (m
6	669.5	23.5	607	13	Q9PS82	Q9ps82 gallus gall
7	669.5	23.5	1014	11	Q921K2	Q921k2 mus musculu
8	653	22.9	607	13	Q9PS81	Q9ps81 xenopus . na
9	636.5	22.3	945	5	Q9N4H4	Q9n4h4 caenorhabdi
10	629	22.0	593	5	Q9TX05	Q9tx05 drosophila
11	598.5	21.0	983	10	Q92P54	Q9zp54 arabidopsis
12	598.5	21.0	1009	10	Q9SJW4	Q9sjw4 arabidopsis
13	578	20.3	969	10	024570	024570 zea mays (m
14	578	20.3	980	10	Q92SV1	Q9zsv1 zea mays (m
15	534	18.7	727	5	Q9XUA5	Q9xua5 caenorhabdi
16	306.5	10.7	2276	5	Q9TXQ1	Q9txq1 caenorhabdi
17	302.5	10.6	815	10	Q9SWB4	Q9swb4 glycine max
18	285.5	10.0	815	10	Q9FK91	Q9fk91 arabidopsis
19	138.5	4.9	1327	4	095271	095271 homo sapien
20	137	4.8	1166	4	Q9H2K2	Q9h2k2 homo sapien
21	137	4.8	1265	4	Q9HAS4	Q9has4 homo sapien
22	134	4.7	1156	16	066878	066878 aquifex aeo
23	127	4.4	363	5	Q9H8R9	Q9h8r9 homo sapien
24	125	4.4	1181	5	Q9XZ37	Q9xz37 drosophila
25	125	4.4	1181	5	Q9VBP3	Q9vbp3 drosophila
26	124.5	4.4	429	11	Q924M2	Q924m2 mus musculu
27	122.5	4.3	848	4	Q9Y2K0	Q9y2k0 homo sapien
28	120.5	4.2	419	4	Q9H9X9	Q9h9x9 homo sapien
29	120	4.2	1219	5	Q9NJ23	Q9nj23 aequipecten
30	120	4.2	1229	5	Q9NJ22	Q9nj22 aequipecten
31	120	4.2	1243	5	Q9NJ21	Q9nj21 aequipecten
32	120	4.2	1253	5	Q9NJ20	Q9nj20 aequipecten
33	120	4.2	1951	11	Q17042	Q17042 aequipecten
34	119.5	4.2	927	11	Q91V71	Q91v71 mus musculu
35	119.5	4.2	946	11	Q91XT8	Q91xt8 mus musculu
36	119.5	4.2	1021	11	Q91XU9	Q91xu9 mus musculu
37	119.5	4.2	1052	11	Q91XT7	Q91xt7 mus musculu
38	119	4.2	2167	10	Q9SSU1	Q9ssu1 chara coral
39	119	4.2	2182	10	Q9LW97	Q9lw97 chara coral
40	117	4.1	261	4	Q9H8F2	Q9h8f2 homo sapien
41	116.5	4.1	1435	3	Q03291	Q03291 saccharomyc
42	116	4.1	429	2	Q9LAX7	Q9lax7 streptococc
43	115.5	4.1	523	11	Q9DA63	Q9da63 mus musculu
44	115.5	4.0	744	4	Q9BXY9	Q9bxy9 homo sapien
45	114.5	4.0	1214	10	Q9FHI5	Q9fhi5 arabidopsis

ALIGNMENTS

```
RESULT 1
Q96CG2

ID   Q96CG2          PRELIMINARY;     PRT;    533 AA.
AC   Q96CG2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 60.1 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014260; AAH14260.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 533 AA;  60089 MW;  6296A0E439CC7767 CRC64;

Query Match             98.9%;  Score 2823;  DB 4;  Length 533;
Best Local Similarity   100.0%;
Matches  533;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    8 MAPKPKPWVQTEGPEKKKGKRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 67
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAPKPKPWVQTEGPEKKKGKRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGT 60

Qy   68 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGGQSKINHFTRLED 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGGQSKINHFTRLED 120

Qy  128 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 180

Qy  188 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQIARGFEALEEEAL 247
```

```
Db  181 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 240     Qy  367 KVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 426
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||            ||:||.|||||||||||||||||:||||||||.||||||||||||||||||||||||||
Qy  248 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA 307     Db  355 KVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPHSGGRVGKGIYFASEN 414
         |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db  241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA 300     Qy  427 SKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDP 486
                                                                                      ||||||||||||||:|||||||||||||||||||||||||.|||||||.|||||||||
Qy  308 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 367     Db  415 SKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPSLKSPPPGFDSVIARGQTEPDP 474
         ||| |||||||||||||||||||||||||||||||||||||||||||::|||||:|||
Db  301 VSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 360     Qy  487 TQDTELELDGQQVVVPGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 540
                                                                                      |||:||||||||||||:|| ||||:||:||||:||||||:||||||||||:|
Qy  368 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGGRVGKGIYFASENS 427     Db  475 AQDIELELDGQPVVVPGGPPVQCPSFKSSSFSQSEYLIYKESQCRLRYLLEIHL 528
         |||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  361 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 420
                                                                              RESULT    3
Qy  428 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 487     Q9TX06
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||            ID   Q9TX06          PRELIMINARY;        PRT;        612 AA.
Db  421 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 480     AC   Q9TX06;
                                                                              DT   01-MAR-2000 (TrEMBLrel. 13, Created)
Qy  488 QDTELELDGQQVVVPGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 540     DT   01-MAR-2000 (TrEMBLrel. 13, Last sequence update)
         ||||||||||||||||||||||||||||||||||||||||||||||||||            DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Db  481 QDTELELDGQQVVVPGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVHL 533     DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
                                                                              OS   Dictyostelium discoideum (slime mold).
RESULT    2                                                                  OC   Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
Q91YR6                                                                       OX   NCBI_TaxID=44689;
ID   Q91YR6          PRELIMINARY;        PRT;        528 AA.              RN   [1]
AC   Q91YR6;                                                                 RP   SEQUENCE FROM N.A.
DT   01-DEC-2001 (TrEMBLrel. 19, Created)                                    RX   MEDLINE=96007847; PubMed=7578427;
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)                      RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)                    RA   Kofler B., Schweiger M., Wagner E.F.;
DE   HYPOTHETICAL 59.4 KDA PROTEIN.                                         RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
OS   Mus musculus (Mouse).                                                   RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      RT   inactivation of the ADPRT gene in the mouse.";
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.     RL   Biochimie 77:444-449(1995).
OX   NCBI_TaxID=10090;                                                       DR   HSSP; P26446; 1A26.
RN   [1]                                                                     DR   InterPro; IPR001290; PARP.
RP   SEQUENCE FROM N.A.                                                      DR   InterPro; IPR004102; PARP_reg.
RA   Strausberg R.;                                                          DR   Pfam; PF00644; PARP; 1.
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.                DR   Pfam; PF02877; PARP_reg; 1.
DR   EMBL; BC014870; AAH14870.1; -.                                          SQ   SEQUENCE   612 AA;  69241 MW;  CB340F7A88FF2364 CRC64;
KW   Hypothetical protein.
SQ   SEQUENCE   528 AA;  59413 MW;  82EFB0C498EB5F74 CRC64;                 Query Match             29.0%;  Score 828;  DB 5;  Length 612;
                                                                            Best Local Similarity   38.4%;  Pred. No. 3.5e-55;
Query Match             78.8%;  Score 2249.5;  DB 11;  Length 528;          Matches  218;  Conservative  79;  Mismatches  197;  Indels  74;  Gaps  19;
Best Local Similarity   80.7%;  Pred. No. 1.6e-164;
Matches  431;  Conservative  37;  Mismatches  59;  Indels  7;  Gaps  4;     Qy    6 LAMAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTARALKAIPAEKRIIRV-------DP 57
                                                                                     |:| | |:|| | ::| | | | |: |:::|| | |::|  ||
Qy    8 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTARALKAIPAEKRIIRVDPTCPLSSNPGT 67     Db   82 LTTRPTKKTKTNTSEAGTK---TAASDLDDSSESE-----DEKNQISVKIKGRAANDP 132
         | ||::| | | ||| | |: : :| | :||| | | |: |:||||||||||||||
Db    1 MAPRKASVQTEG--SKKQRQGTEEEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58     Qy   58 TCPLSSNPGTQOYYED-----YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVG 112
                                                                                     |:| | | |:| |  |:| || |: | : :|| : |  |
Qy   68 QVYEDYNCTLNQTNIENNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLED 127     Db  133 HFPDSR--XHVYENGKDVYDATLNQTEISQNNNKXYIIQLLEADDGSSYSVWNRWGREG 189
         |::|| ::|:|:| | |||||||||||||| :::|||||||||||: :||: |  ||
Db   59 QVHEDYDCTLNQTNIGNGNNNKFYIIQLLEEGSRFF-CWNRWGRVGEVGQSKMNHFTCLED 117     Qy  113 EVGQSKINHFTR--LEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQE 170
                                                                                     | || || |   |  :|| | | || |  | :: |:| | | : |:| | |
Qy  128 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 187     Db  190 LKGQSSRKDFGKGGLNQAISLFCSKFYEKTKNTFTDRANFKKVAGKYDMIELDYSTDS-- 247
         ||||::|||: ::||: :||||||||||||||:|||||||| :|:  | | | |  |||
Db  118 AKKDFKKKFWEKTKNKWEERDRFVAQPNKYTLIEVQGEASQEAVVKVDSGPVRTV---V 174     Qy  171 AVVKVDRG--------PVRTVTKRV----QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKK 222
                                                                                     :| |   | |: ::|   :|| || | |: |  | | :: |:| | |:| |
Qy  188 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 247     Db  248 ---KPKNGASTTATTTTTKKVVEHKKECSLDERVQELVKLIFDVKMERTMTEAKYDLKK 304
         |||||||||:||||||||||||||||||||||||||||||||:|||||||||||||||
Db  175 KPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTKQQIARGFEALEALEEAM 234     Qy  223 MPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINS 282
                                                                                     ||||||:|||| | :|||  | | | |  |:| :: | | | | |: |::
Qy  248 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA 307     Db  305 MPLGKLSKNQITKGYLVLKQIEDVMGGKS--GESLSTLSSRFYTIIPHAFGMSVPPVINT 362
         ||||:||||||||| ::| |||||||||| |::| | | | |:: |  | | | | :
Db  235 KNPTGDGQSLEELSSLSCFYTVIPHNFGRSRPPIINSPDVLQAKKDMLLVLADIELVQTLQA 294     Qy  283 PELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPE 342
                                                                                     :|: | | | |:: | | |  |:| | :|  | | |:| |: | |
Qy  308 V-SEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIW 366     Db  363 NQMLIEKMNMLQNLADIEIATNIIKDSESDES------NILELHYAKLKTIQPLDENSCE 417
         |  |:|:| |: | | | |:| | | |  | |:: | |  | | | : | | |
Db  295 APGEEEEKVEVPHPLDRDYQLLRCQLQLLDSGESEYKAIQTYLKQTGNSYRCPNLRHVW 354     Qy  343 YKVIQTYLEQTGSNHRCPTLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAIL 402
                                                                                     | : | | | : : | | | : | | | | | | : || |:|| ||:|| | :
                                                                              Db  418 YKNILLYVKNTYQGGKPTIVNIFKIDRDGEADRYKTKKHLGNRKLWHGSRLTNYASII 477
```

```
Qy   403 TSGLRIMPH---SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY-----MFLGEVAL 453
              : !!!! !    !! ! !!!!!! ! !!!!!    !!   : !!:!!!
Db   478 SQLRIAPPEAPVSGYRFGKGVYFADCMSLSANY------CRTVGFLIDFCMLLGDVAL 530
Qy   454 GREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT---QDTELELDGQQVVVPQGQPVPCP 510
             ! !!  :: !!   !    !     !!  !!   : !! !! :!
Db   531 GKTADLARDT-YMEKPQPNSHSTWALGTVEPDPKVFFQDTE-----GCITPYGQMIPSQ 583
Qy   511 EFSSSTFSQSEYLIYQESQCRLRYLLEV 538
              : :!:! : !!!!!!:!!!!:
Db   584 HKGVSCY-EHQYVVVDVAQVHLKYLLQL 610
```

RESULT 4
081294
ID   081294;   PRELIMINARY;   PRT;   635 AA.
AC   081294;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE).
GN   T14P8.19 OR AT4G02390.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   WASHU;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Kalick1 J., Elliott G., Cloud J.;
RT   "The sequence of A. thaliana T14P8.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Waterston R.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Lamar B., StoneKing T., Stumpf J., Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF069298; AAC19283.1; -.
DR   EMBL; AL161494; CAB80732.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
KW   Transferase.
SQ   SEQUENCE   635 AA; 72017 MW; E3F1CBE4D367A377 CRC64;

Query Match   24.3%; Score 694; DB 10; Length 635;
Best Local Similarity 34.2%; Pred. No. 7.4e-45;
Matches 185; Conservative 97; Mismatches 221; Indels 38; Gaps 18;

```
Qy    18 TEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRV-DPTCPLSSNPGTQVYED----72
              :  :  : !  :!:   !!   !  ! !  !  ! !!! !  :
```

```
Db   107 SNAPVKSSNDEAEDDNGFEEEKKEEKIVTATKKGAAVLDQWIPDEIKSQYHVLQRGDDV 166
             !   :!!!!!!  :!  ::!:!   !!!!
Qy    73 YNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKIN-HFTRLEDAKK 130
             !: !!!!!:  !!!! ! :! ::: !! !! ! !!!!!  ! ! !  - ! : !
Db   167 YDAILNQTNVRDNNNKFFVLQVLESDSKKTYMYTRWGRVGVKGQSKLDGPYDSWDRAIE 226
Qy   131 DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQP- 189
             ! !! !!!! :!!!!  ::!!!!!  :  :! ::! ::! : :  !!       !:!
Db   227 IFTNKFNDKTKNYWSDRKEFIPHPKSYTWLEMDYGKEENDSPVNND---IPSSSSEVKPE 283
Qy   190 -CSLDPATQKLTNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALK 248
             : ! !  !   : !:! !: !!   !!: !!  !!!!:!: !:::!:!! !:: ! :!
Db   284 QSKLDTRVAKFISLICNVSMMAQHMEIGYNANKLPLGKISKSTISKGYEVLKRISEVI- 342
Qy   249 GPTDGGQSLEELSSHFYTVIPHNFGHSQPPP-INSPELLQAKKDMLLVLADIELAQALQ 306
             !!!!!!!  : !!!!!!!!!   !  ! !!! !  : !:!!!::!:! !:! !!!!::!!
Db   343 -DRYDRTRLEELSGEFYTVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEIELATKLL 401
Qy   307 AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ- 363
             :!    !  !! !!       :! !!!!!!!!!: !  :  !!!  !
Db   402 SVDPGLQ-----DDPLYHYQQLNCGLTPVGNDSEEFSMVANYMENTHAKTHSGYTVEIA 456
Qy   364 HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGKG 419
              :  :: :  !!:!!   !!!:!!!! !!!!! !  !:! !  : !!!!   !!!!!!!!
Db   457 QLFRASRAVEADRFQQFSSSKNRMLLWHGSRLTNWAGIISQGLRIAPPEAPVTGYMFGKG 516
Qy   420 IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 479
             :!!!!!!!!!!!   ! !!!!!   !  !!!!!!!!!
Db   517 VYFADMFSKSANYCYA-NTGAND-GVLLLCEVALGDMNELLYSDYNADNLPPGKLSTKGV 574
Qy   480 GHTEPDPTQDTELELDGQQVVVPQGQPV-PCPEFSSSTFSQSEYLIYQESQCRLRYLLE 537
             ! !:!:!     !!!!!!!!!    !!!  :   :! !!!! !! :! !!!:!!!!!!!
Db   575 GKTAPNPSEAQTLE-DG--VVVPLGKPVERSC---SKGMLLYNEYIVYNVEQIKMRYVIQ 628

Qy   538 V 538
             !
Db   629 V 629
```

RESULT 5
O50017
ID   O50017   PRELIMINARY;   PRT;   653 AA.
AC   O50017;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE.
GN   PARP.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babychuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA   O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT   "Higher plants possess two poly(ADP-ribose) polymerases.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ222588; CAA10888.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF02037; SAP; 2.
DR   SMART; SM00513; SAP; 2.
SQ   SEQUENCE   653 AA; 72995 MW; 5FD01923C4ABCD1D CRC64;

Query Match   23.5%; Score 670.5; DB 10; Length 653;
Best Local Similarity 35.4%; Pred. No. 4.9e-43;

```
Matches  170;  Conservative  93;  Mismatches 186;  Indels  31;  Gaps 14;

Qy   71 EDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEVGQSKINHFTRLED-A 128
            |:| |||||||||:||||||||:||: |   |   | ||
Db  188 EIYDATLNQTNVGDNNNKFYIIQVLESDAGGSFMVYNRWGRVGVRGQDKIJHGPSPTRDQA 247

Qy  129 KKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ 188
            ||| || |||||||||||||||:|:  || ::| ::   ::|
Db  248 IYEFEGKFHNKTNNHWSDRKNFKCYAKKYTWLEMDYGETEKE---SITDQIK 300

Qy  189 PCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKILSKQQIARGFEALEALEEALK 248
            ||:| || |||||||||||||||||:||  :| : |
Db  301 ETKLETRIAQFISLICNISMKQRMVEIGYNAEKLPLGKLRKATILKGYHVLKRISDVIS 360

Qy  249 GPTDGGQSLEELSSHFYTVIPHNFGHSQPPP--INSPELLQAKKDMLLVLADIELAQALQ 306
            ||| ||  ||:| :|:  |:| |||| ||:
Db  361 --KADRRHLEQLTGEFYTVIPHDFGFRKMREFIIDTPQKLKAKLEMVEALGEIEIATKL- 417

Qy  307 AVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ- 363
            || |:||||  || ||| |||||| |:|||||| : |
Db  418 --LEDDSDQDDPLYARYKQLHCDFTPLEADSDEYSMIKSYLRNTHGKTHSGYTVDIV 473

Qy  364 HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH---SGGRVGKG 419
            ||||||:||||||||||||||:||||||| |  |
Db  474 QIFKVSRHGETERFQKFASTRNRMLLWHGSRLSNWAGILSQGLRIAPPEAPVTGYMFGKG 533

Qy  420 IYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 479
            ||:|  |      ||||||| || || |
Db  534 VYFADMFSKSANYCYASE--ACRSGVLLLCEVALGDMNELLNADYDANNLPKGKLRSKGV 591

Qy  480 GHTEPDPTQDTELELDGQQVVPGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEVH 539
            ||  |  ||:  | |:|| :|| | ||
Db  592 GQTAPNMV-ESKVADDG--VVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMRYVLHVN 647
```

RESULT 6
Q9PS82
```
ID   Q9PS82      PRELIMINARY;      PRT;   607 AA.
AC   Q9PS82;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007827; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE  607 AA;  68033 MW;  75F6EE1D30D8F402 CRC64;
```

```
Query Match             23.5%;  Score 669.5;  DB 13;  Pred. No. 5.3e-43;
Best Local Similarity   33.2%;
Matches  186;  Conservative 100;  Mismatches 205;  Indels 69;  Gaps 24;

Qy   14 PW---------VQTEGPE--------KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPT 58
```

```
Db   73 PWGAEVKTEHQEVAVDGKCSKPANMKSAGKVKEEQGPSKSEKKMKLTV--KGGAAVDPD 129
            ||     :|||            :          |  : |||
Qy   59 CPLSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQD-SNRFFTCWNRWGRVG 112
            |           ::| ||   |   |   | ||
Db  130 SGLEDS--AHVFEKGGKIFSATLGLVDIVKGTNSYYKLQLLEDDRESRYWV-FRSWGRVG 186

Qy  113 EV-GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQE 170
            :| |  |      ||| ||||||||||||||||||||||:|| ||| ||:|:
Db  187 TVIGSNKLEQMPSKEDAVEHFLNLYEEKTGNSWHSK-NFTKYPKKFYPLEIDYGQDEEAV 245

Qy  171 AVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSK 230
            |: |  :|: ||  |||||||:| ||||:||||||||||||:|||||||||||||||||
Db  246 RKLTVSAG--TKSKLAKP-----IQDLIKMIFDVESMKKAMVEFEIDLQKMPLGKLSK 296

Qy  231 QQIARGFEALEALEEALKGPTDGGQSLE--ELSSHFYTVIPHNFGHSQPPPINSPELLQA 288
             :|:   :::   :: :| ||:||:|    ||||:|||:|||:||  ||| || ||::|
Db  297 RQIQSAYSILNEVQQAV--SDGGSESQILDLSNRFYTLIPHDFGMKKPPLLSNLEYIQA 353

Qy  289 KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQT 348
            | :||  | ||: |  |               | |  ||     ::|   |  : ::|
Db  354 KVQMLDNLLDIEVAYSLLRGGNEDGDKD---PIDINYEKLRTDIKVVDKDSEEAAKIIKQ 409

Qy  349 YLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSG 405
            | :    ::| |  :|   |:|  |  |||  | | ::: :|:||||| : ::||:|:|
Db  410 YVKNTHAATHNAYDLKVVEIFRIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGILSQG 469

Qy  406 LRIMPH----SGGRVGKGIYFASENSKSAGYVIGMKC--GAHHVGYMFLGEVALGREHH 458
            |||    |   ||:|||||||    :|||:        | | |  ||| ||||||
Db  470 LRIAPPEAPVTGYMFGKGIYFRDMVSKSANY----CHTSQADPIGLILLGEVALGNMYE 524

Qy  459 INTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVVCPEFSSSTFS 518
            :|   |  || ||| :  | |  | || |   |||  ::    | | |  ||    ::
Db  525 LKNASHITKL-PKGKHSVKGLGKTAPDPTATT--TLDG--VEVPLGNGI-STGINDTCLL 578

Qy  519 QSEYLIYQESQCRLRYLLEV 538
            :||:|||  :|||||:||::
Db  579 YNEYIVYDVAQVNLKYLLKL 598
```

RESULT 7
Q921K2
```
ID   Q921K2      PRELIMINARY;      PRT;   1014 AA.
AC   Q921K2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE)
DE   POLYMERASE).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC012041; AAH12041.1; -.
KW   Transferase.
SQ   SEQUENCE  1014 AA;  112721 MW;  5DCE68E4CB3F46EB CRC64;
```

```
Query Match             23.5%;  Score 669.5;  DB 11;  Length 1014;
Best Local Similarity   33.5%;  Pred. No. 1.1e-42;
Matches  186;  Conservative  99;  Mismatches 203;  Indels  67;  Gaps 22;

Qy    8 MAPKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRI-----IRVDPTCP 60
            :||     ::|                                :    ||||:|
Db  494 VAPKGK-----SAAPSKKSKGAVKEEGVNKS---------EKRMKLTLKGGAAVDPDSG 538

Qy   61 LSSNPGTQVYED---YNCTLNQTNIENNNKFYIIQLLQDSNRF-FTCWNRWGRVGEV- 114
            | |           :| ||   |   | ||
Db  539 LEHS--AHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKESRYWIFRSWGRVGTVI 596
```

```
Qy   115 GQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVV 173
             |:||    ||:     |  :    |  :||||||||||| :  ||:   ::|   |||
Db   597 GSNKLEQMPSKEDAVEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVK 653

Qy   174 KVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQI 233
          | :|:|  |: || | | || : : || :  ||  ||||||||||
Db   654 KLTVKP-GTKSKLPKP------VQELVGMIFDVESMKKALVEYEIDLQKMPLGKLSRRQI 706

Qy   234 ARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDML 293
           ..: ||  :| | |.||: : :   |||:|||:||| |||:  ||| | ||  ||:||
Db   707 QAAYSILSEVQQAVSQGSSESQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEML 765

Qy   294 LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 353
          : |||::  : |      :|   : || || ::||  ||  ||| |     |||    |
Db   766 DNLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEVIRKVVKNT 821

Qy   354 -GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMP 410
                 ||::|   ||:|  |    ||: || :|||| ||||||:|  | :|| ||||| |
Db   822 HATTHNAYDLEVIDIFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGILSQGLRIAP 881

Qy   411 H----SGGRVGKGIYFASENSKSAGYVIGMKCGAHH---VGYMFLGEVALGREHHINTDN 463
          .     || :|||||||| ||||| ||||||  |       |:| |||||| |  | .:
Db   882 PEAPVTGYMFGKGIYFADMVSKSANY----CHTSQGDPIGLILLGEVALGNMYELKHAS 936

Qy   464 PSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYL 523
          . |.   ||  ||     |.||| :   :|||      |   ||||  | ||  ||:|:
Db   937 -HISKLPKGKHSVKGLGKTTPDPSASITLE---GVEVPLGTGIP-SGVNDICLLYNEYI 990

Qy   524 IYQESQCRLRYLLEV 538
          :|    ||:|:||::
Db   991 VYDIAQVNLKYLLKL 1005
```

```
Db    89 PSSGPVAGKSSGKVKEEKGSNKSEKKMKLTVKGGAAIDPDSEL---EDSCHVLETGG--- 142
          .       ::  |  ||| ||| |  :|:  .       |.:|    |||  || :
Qy    70 YEDYNCTLNQTNIENNNKFYIIQLLQ-DSNRFFTCWNRWGRVGEV-GQSKINHFTRLED 127

Db   143 -KIFSATLGLVDITRGTNSYYKLQLIEHDRDSRYWVFRSWGRVGTVIGSKKLEEMSSKED 201
             | :     |  : |  |  ::|:|   | |  |  .  | | |.   ||
Qy   128 AKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRV 187

Db   202 AIEHFLNLYQDKTGNAW-HSPNFTKYPKKFYPLEIDYGQE-EDVKKLSVG-AGTKSKLA 258
           | |       |  |   | .|    ||||| || | |   :||||  |.  |||||
Qy   188 QPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEAL 247

Db   259 KP------VQELIKLIFDVESMKKAMVEFEIDLQKMPLGKLSKRQIQSAYSILSQVQQAV 312
          ||       |||:|||||||:|||| ||||| |||||||||||:|| |   ::| |||
Qy   248 KGPTDGGQSLEE------LSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELA 302

Db   313 ------SESLSEARLLDLSNQFYTLIPHDFGMKKPPLLNNLEYIQAKVQMLDNLLDIEVA 366
               |||  :| || ||||:||:||| ||: | ||| |     ||| |:|  |:||:|
Qy   303 QAL--QAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRC 359

Db   367 YSLLRGGADDGEK------DPIDVKYEKIKTDIKVVAKDSEESRIICDYVKNTHADTHNA 420
          ||| |||:||||        |||| |||:||||||| |||||: :||| ||||||||||
Qy   360 PTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SG 413

Db   421 YDLEVLEIFKIDRRGEYQRYKPFKQLHNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTG 480
             |  :|  | .  | |||| | : |||: ||||| :  |||:|| |||||  |  | |
Qy   414 GRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGF 473

Db   481 YMFGKGIYFADMVSKSANYCHAMP--GSPIGLILLGEVALGNMHELKAASQITKL-PKGK 537
          : |||||||| ||||||| |  |    |.|| |||||||||  |  | .:| |||  ||
Qy   474 DSVIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLR 533

Db   538 HSVKGLGRTAPDPS-ATVQLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLK 592
```

```
Qy   534 YLLEV 538
          |||::
Db   593 YLLKL 597
```

```
RESULT 9
Q9N4H4
ID   Q9N4H4;      PRELIMINARY;      PRT;   945 AA.
AC   Q9N4H4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 108.0 KDA PROTEIN.
GN   Y71F9AL.18
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditina;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Bradshaw-Cordum H., Scott K., Graves T.;
RT   "The sequence of C. elegans cosmid Y71F9AL.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024200; AAF36011.1; -.
```

```
RESULT 8
Q9PS81
ID   Q9PS81;      PRELIMINARY;      PRT;   607 AA.
AC   Q9PS81;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+ PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Xenopus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae.
OX   NCBI_TaxID=8353;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";
RL   Biochimie 77:444-449(1995).
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   PROSITE; PS50172; BRCT; 1.
SQ   SEQUENCE   607 AA;  67496 MW;  54CDEBBE22079886 CRC64;

Query Match            22.9%;  Score 653;  DB 13;  Length 607;
Best Local Similarity   32.5%;
  Matches  177;  Conservative  102;  Mismatches  214;  Indels   52;  Gaps   20;

Qy    10 PKPKPWVQTEGPEKKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQV 69
          | |  |     | |  || :  |:||| |  :|  || : : | |||||  | |  |
```

Pred. No. 9.8e-42;

```
DR   HSSP; P26446; 1A26.                          RL   Biochimie 77:444-449(1995).
DR   InterPro; IPR001290; PARP.                   DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001510; Znf-PARP.               DR   InterPro; IPR001357; BRCT.
DR   Pfam; PF00644; PARP; 1.                      DR   InterPro; IPR001290; PARP.
DR   Pfam; PF00645; zf-PARP; 1.                   DR   InterPro; IPR004102; PARP_reg.
DR   ProDom; PD004675; Znf-PARP; 1.               DR   Pfam; PF00533; BRCT; 1.
DR   PROSITE; PS50064; PARP_2N_FINGER_2; 1.       DR   Pfam; PF00644; PARP; 1.
KW   Hypothetical protein.                        DR   Pfam; PF02877; PARP_reg; 1.
SQ   SEQUENCE  945 AA;  108006 MW;  1D0A62C954BC6AD9 CRC64;   DR   PROSITE; PS50172; BRCT; 1.
                                                  SQ   SEQUENCE  593 AA;  68018 MW;  D9BA37E38B8E7CCD CRC64;

Query Match               22.3%;  Score 636.5;  DB 5;  Length 945;   Query Match               22.0%;  Score 629;  DB 5;  Length 593;
Best Local Similarity  34.0%;  Pred. No. 3.5e-40;                    Best Local Similarity  32.1%;  Pred. No. 6.6e-40;
Matches 167; Conservative 84; Mismatches 189; Indels 51; Gaps 15;    Matches 167; Conservative 100; Mismatches 196; Indels 58; Gaps 19;

Qy    73  YNCTLNQTNIENNNKFYIIQLLQDSNR-FFTCWNRWGRVG-EVGQSKINHFTRLEDAKK  130   Qy    44  KAIPAEKRI-----IRVDPTCPLSSNPGTQVYED----YNCTLNQTNIENNNNKFYIIQL   94
          | |:| || :|| |||:|| : :  ||||| :  ::: ||  :: :: :| :|:|           |  :|| ||       | :| | : |  ||  :  :  :| :  : :|:|: |:|: |
Db   477  YQATLSFTDLTQNRKNSYYKIQLLKDDQRENYYVFRSWGRVGTEVGGNKHESYSNSNEAIL  536   Db    99  KSMPVSRTFKVKDGLAVDPDSGLEDI--AHYVDSNNKYSVVLGLTDIQRNKNSYYKVQL  156

Qy   131  DFEKKFREKTKNNWAERDHFVSHPGKYTLLEVQAEDEAQEAVVKVDRGPVRTVTKRVQPC  190   Qy    95  LQ-DSNRFFTCWNRWGRVG-EVGQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVS  152
          || | ||| || |  |:| : |   |:| | :|| :|| :| : :  : ||||         ||  ||::|||| ||:|||| ||||:||||||||| |||:|:|:::| ::|
Db   537  KFQDVFHEKTKNDWIYRKHFRKMPCMFSYVETDYSEFAQITDTEITPG----SKTLLPK  591   Db   157  LKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSESAKRNFEIYADKTGNEYEQRDNFVK  216

Qy   191  SLDPATQKLITNIFSKEMPKNTMALMLDLDVKKMPLGKLSKQQIARGFEALEALEEAL-KG  249   Qy   153  HPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNT  212
          |:      |||  :: :| :       ::|||  | ||   || :| : ||  |  ||          :: |||:|:  :: | | :  :  :|||
Db   592  SV-----KEVVMSIFDVENNKSALKSFEMDVNKMPLGRLSHNQINLAFEVLNDISDLLVKL  647   Db   217  RTGRMYPIEIQYDDD--QKLVKHE------SHFFTSKLEISVQNLIKLIFDIDSMNKT  266

Qy   250  PTDGGQSLEELSSHFYTVIPHNFGHSQQPPINSPELLQAKKDMLLVLADIELAQALQAVS  309   Qy   213  MALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNF  272
          | : | |  |:|||||| || |||::  | || :  |  ||| :: |  :::: :  :|:::          |: ||  | |||::  :|   || :| :| :  :: :|:||||:||||||
Db   648  PIDASRIL-DFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNMLNALLDIKFAYDQISGG  706   Db   267  LMEFHIDMDKMPLGKLSAHQIQSAYRVKEIYNVLECGSNTAK-LIDATNRFYTLIPHNF  325

Qy   310  EQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-GSNHRCPT-LQHIWK  367   Qy   273  GHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQ  332
          | :  ||:|| | : |  :|  :|  |: | |||::| :| ||  |||| | |  | | |          |::::||| ||  |:  |:: ||  :||||| ::
Db   707  DVPASTSLGIDPVDINYQKLKCIMEPLQQGCDDWNMIHQYLKNTHGATHDLKVELIDILK  766   Db   326  GVQLPFLIETHQQIEDLRQMLDSLAEIEVAYSI----IKSEDVSDACNPLDNHYAQIKTQ  381

Qy   368  VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGKGIYFA  423   Qy   333  LQLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLL  389
          :: :|: ||| |||||| ||:|:||: | :| | ||         :|||| ||||||          | | |: | |:|:| :| ||  ::|:  |  : || ||| ||||| ||||
Db   767  LNRDNESSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPPEAPVSGYMFGKGVYFA  824   Db   382  LVALDKNSEEFSILSQYVKNTHASTHKSYDLKIVDVFKVSRQGEARREKPFKLHNRKLL  441

Qy   424  SENSKSAGYVIGMKCGAHHVGYMFLGEVALGR-EHHINTDNPSLKSPPPGFDSVIARG-  480   Qy   390  WHGTNMAVVAAILTSGLRI----MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGY  445
            :  ||| |:   |||||||| ||| | | ||   :: :   |||  |  || :: ||          | | | :| :||||: :| ::|| | | || | |::|||   |||||||
Db   825  DMFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSRQTLPAGFQSVQGLGRQ  881   Db   442  WHGSRLTNFVGILSHGLRIAPPEAPPTGYMFGKGIYFADMVSKSANYCCTSQ--QNSTGL  499

Qy   481  ------HTEPD----PTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQ  529   Qy   446  MFLGEVALG-------REHHIN-TDNPSLKSPPPGFDSVIARGHTEPDTQDTELELDGQ  497
          |||: :  |  || | ||: |||:| |:|||||   |: |:| |:| |          |:| ||:|             |:||  |||::::|: |:|: :
Db   882  CPREIGSYNHPDGYTIPLGLTYMQLQGKQDV------------DYHLLYNEFIYDVDQ  928   Db   500  MLLSEVALGDMMECTSAKYINKLSNNK-----HSCFGRGRTMPDPTK-SYIRSDG-  548

Qy   530  CRLRYLLEVHL  540                                              Qy   498  QVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLLEV  538
          :|:|| || :                                                            :||  | |:: :||::|| |  ::|:| ||: ||
Db   929  IQLKYLVRVKM  939                                             Db   549  -VEIPYGETI-TDEHLKSSLLYNEYIVYDVAQVNIQYLFRM  587


RESULT 10
Q9TX05
ID   Q9TX05       PRELIMINARY;       PRT;    593 AA.
AC   Q9TX05;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NAD+PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS   Drosophila sp. (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7242;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96007847; PubMed=7578427;
RA   Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA   Kofler B., Schweiger M., Wagner E.F.;
RT   "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT   ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT   inactivation of the ADPRT gene in the mouse.";


RESULT 11
Q9ZP54
ID   Q9ZP54       PRELIMINARY;       PRT;    983 AA.
AC   Q9ZP54;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).
GN   PARP-1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LANDSBERG ERECTA;
RA   Doucet-chabeaud G., Kazmaier M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ131705; CAA10482.1; -.
```

```
DR   HSSP; P26446; 1A26.
KW   Transferase; Glycosyltransferase; NAD.
FT   CHAIN         2      983       POLY (ADP-RIBOSE) POLYMERASE.
SQ   SEQUENCE   983 AA;  111232 MW;  46E12A8EF1B6F4F CRC64;
```

```
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005593; AAD20677.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001357;  BRCT.
DR   InterPro; IPR001290;  PARP.
DR   InterPro; IPR004102;  PARP_reg.
DR   InterPro; IPR001510;  znf-PARP.
DR   Pfam; PF00533;  BRCT; 1.
DR   Pfam; PF00644;  PARP; 1.
DR   Pfam; PF02877;  PARP_reg; 1.
DR   Pfam; PF00645;  zf-PARP; 2.
DR   ProDom; PD004675; Znf-PARP; 2.
DR   SMART; SM00292; BRCT; 1.
DR   PROSITE; PS50180;  BRCT; 1.
DR   PROSITE; PS50064;  PARP_ZN_FINGER_2; 2.
SQ   SEQUENCE   1009 AA;  114133 MW;  CDE6E41CC2A3A2DB CRC64;
```

```
Query Match             21.0%;  Score 598.5;  DB 10;  Length 1009;
Best Local Similarity   31.1%;  Pred. No. 3.2e-37;
  Matches  171;  Conservative 101;  Mismatches 199;  Indels  79;  Gaps  23;
```

```
Qy     24  KKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQ----VYED---YNC  75
               |||:| ||   :   :| | :  |:|::|| :| :          :          ||
Db    498  KKQRKLPFDKYKIEDTSESLVTVKVKGR------SAVHEASGLQEHCHILEDGNSIYNT  550

Qy     76  TLNQTNIENNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD  131
           ||::: :  |||| ::| || :    :|    :||||   ::|||:|| :  : :|
Db    551  TLSMSDLSTGINSYYILQIIQE-DKGSDCYVFRKWGRVGNEKIGGNKVEEMSK-SDAVHE  608

Qy    132  FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ  188
           |::| || |      : | | |   :: |:     :: |||        ||  | | :
Db    609  FKRLFLEKTGNTWESWEQKTNFQKPGKFLPLD--------------IDYGVNKQVAKK-E  654

Qy    189  P----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE  244
                 |:|||  |    ||::|   :  ::|:|| :|||:|||| :||::::|   : ||:
Db    655  PFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ  714

Qy    245  EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIE  300
           ||      ||||  || || :|:||:  :::|:| ||:      :|||::  |:||||:
Db    715  EAL---KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIE  770

Qy    301  LAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP  360
           : |:| | |::||  ::   :| :   | || :|| :| |   :  : : :|:::::|||
Db    771  IASRI--VGFDVDSTES---LDDKYKKLHCDISPLPHDSEDYRLIEKYLNTT----HAP  820

Qy    361  T-------LQHIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRI----  408
           |        ||||| : ||| || :  ||:|||:|||||| |  |:|  | ::||
Db    821  THTEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPE  880

Qy    409  MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS  468
           : : |||||||||| |||  :|||  |  |  | || ||| |:::  :| |:   |  |:
Db    881  APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGEIHEL-TKAKYMDK  937

Qy    469  PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES  528
           || | | :   |:: || |  |   ::|  | || |  || |:| ||::::  | :::
Db    938  PPRGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEYIVYDTA  993

Qy    529  QCRLRYLLEV  538
           | | |:|:
Db    994  QVKLQFLLKV  1003
```

```
RESULT  12
ID   Q9SJW4
AC   Q9SJW4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE POLY (ADP-RIBOSE) POLYMERASE.
GN   AT2G31320.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
```

```
Query Match             21.0%;  Score 598.5;  DB 10;  Length 983;
Best Local Similarity   31.1%;  Pred. No. 3.1e-37;
  Matches  171;  Conservative 101;  Mismatches 199;  Indels  79;  Gaps  23;
```

```
Qy     24  KKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQ----VYED---YNC  75
               |||:| ||   :   :| | :  |:|::|| :| :          :          ||
Db    472  KKQRKLPFDKYKIEDTSESLVTVKVKGR------SAVHEASGLQEHCHILEDGNSIYNT  524

Qy     76  TLNQTNIENNNKFYIIQLLQDSNRFFTCW--NRWGRVG--EVGQSKINHFTRLEDAKKD  131
           ||::: :  |||| ::| || :    :|    :||||   ::|||:|| :  : :|
Db    525  TLSMSDLSTGINSYYILQIIQE-DKGSDCYVFRKWGRVGNEKIGGNKVEEMSK-SDAVHE  582

Qy    132  FEKKFREKTKN---NWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQ  188
           |::| || |      : | | |   :: |:     :: |||        ||  | | :
Db    583  FKRLFLEKTGNTWESWEQKTNFQKPGKFLPLD--------------IDYGVNKQVAKK-E  628

Qy    189  P----CSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE  244
                 |:|||  |    ||::|   :  ::|:|| :|||:|||| :||::::|   : ||:
Db    629  PFQTSSNLAPSLIELMKMLFDVETYRSAMMEFEINMSEMPLGKLSKHNIQKGFEALTEIQ  688

Qy    245  EAL----KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIE  300
           ||      ||||  || || :|:||:  :::|:| ||:      :|||::  |:||||:
Db    689  RLLTESDPQPTMKESLLVDASNRFFTMIP---SIHPHIIRDEDDFKSKVKMLEALQDIE  744

Qy    301  LAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCP  360
           : |:| | |::||  ::   :| :   | || :|| :| |   :  : : :|:::::|||
Db    745  IASRI--VGFDVDSTES---LDDKYKKLHCDISPLPHDSEDYRLIEKYLNTT----HAP  794

Qy    361  T-------LQHIWKVNQEGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRI----  408
           |        ||||| : ||| || :  ||:|||:|||||| |  |:|  | ::||
Db    795  THTEWSLELEEVFALEREGEFDKYAPHREKLGNKMLLWHGSRLTNFVGILNQGLRIAPPE  854

Qy    409  MPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKS  468
           : : |||||||||| |||  :|||  |  |  | || ||| |:::  :| |:   |  |:
Db    855  APATGYMFGKGIYFADLVSKSAQYC--YTCKKNPVGLMLLSEVALGEIHEL-TKAKYMDK  911

Qy    469  PPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQES  528
           || | | :   |:: || |  |   ::|  | || |  || |:| ||::::  | :::
Db    912  PPRGKHSTKGLGKKVP---QDSEFAKWRGDVTVPCGKPV-SSKVKASELMYNEYIVYDTA  967

Qy    529  QCRLRYLLEV  538
           | | |:|:
Db    968  QVKLQFLLKV  977
```

```
RESULT  12
Q9SJW4
ID   Q9SJW4          PRELIMINARY;     PRT;   1009 AA.
```

```
RESULT 13
ID      O24570          PRELIMINARY;     PRT;     969 AA.
AC      O24570;
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      POLY(ADP-RIBOSE) POLYMERASE.
OS      Zea mays (Maize).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC      Panicoideae; Andropogoneae; Zea.
OX      NCBI_TaxID=4577;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Babiychuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA      O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT      "Higher plants possess two poly(ADP-ribose) polymerases.";
RL      Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR      EMBL; AJ222589; CAA10889.1; -.
DR      HSSP; P26446; 1A26.
DR      InterPro; IPR001357; BRCT.
DR      InterPro; IPR001290; PARP.
DR      InterPro; IPR004102; PARP_reg.
DR      InterPro; IPR003034; SAP.
DR      InterPro; IPR001510; Znf-PARP.
DR      Pfam; PF00533; BRCT; 1.
DR      Pfam; PF00644; PARP; 1.
DR      Pfam; PF02877; PARP_reg; 1.
DR      Pfam; PF00645; zf-PARP; 2.
DR      ProDom; PD004675; Znf-PARP; 2.
DR      SMART; SM00292; BRCT; 1.
DR      PROSITE; PS50172; BRCT; 1.
DR      PROSITE; PS50064; PARP_zN_FINGER_2; 2.
SQ      SEQUENCE  969 AA;  109128 MW;  EB23AC62EEC14009 CRC64;

Query Match             20.3%;  Score 578;  DB 10;  Pred. No. 1.1e-35;
Best Local Similarity   30.6%;
Matches  166;  Conservative  96;  Mismatches  195;  Indels  86;  Gaps  21;

Qy          23 KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNI 82
                 | ||| || :|  | ||     |:  || ::::
Db         480 KVKGRSAVHESSGLQDTAHILE---------------------DGKSI---YNATLNMSDL 516

Qy          83 ENNNNKFYIIQLL-QDSNRFFTCWNRWGRVG--EVGQSKINHFTRLEDAKKDFEKKFREK 139
                 :  |||| :|| :|:||| | :  ||||| :  |||||||  | || :|  | ||
Db         517 ALGVNSYYVLQIIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEK 575

Qy         140 TKNNWAE---RDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPAT 196
                 | :| |      ||| :|:|| :| |   :  ||    |:    |  |:|   |
Db         576 TGNSWEAWECKTNFRKQPGRFYPLDVD-------YGVKKAPKRKDISEMK-SSLAPQL 625

Qy         197 QKLITNIFSKEMPKNTMALMDLDVKKMPLGKLSKQOIARGFEALEALEEALKGPTDGGQS 256
                 || |   |: | :   || :|| |||    | |  |:  ||| | | || || |
Db         626 LELMKMLFNVETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALA 685

Qy         257 LEE-----LSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAV-SE 310
                 |:      | ||| | |||  | || ||: || :| ||    :| ||:  |   :|
Db         686 VRESLIVAASNRFFTLIP----SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSD 741

Qy         311 QEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT------LQ 363
                  ::|   |   | |   :| ||    |:   |  :|:  |:|: | ||  :      |
Db         742 SDES------LDDKYMKLHCDITPLAHDSEDYKLIEQYL----LNTHAPTHKDWSLELE 790

Qy         364 HIWKVNQREGEEDRFQAH-SKLGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGK 418
                 || :     ::| :: |: |||  |::||   | |::||   |  |  |     |||| |
Db         791 EVFSLRDGELNKYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAFPEAPVTGYMFGK 850

Qy         419 GIYFASENSKSAGYVIGMKCGA---HHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDS 475
                 |||| ||| ||  | | :       ::|:  | :| |||      : :|    |||: |
Db         851 GLYFADLVSKSAQY-----CYVDRNNPVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS 904
```
```
Qy         476 VIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYL 535
                 :|:               :|||       |||    ||||||:|:||| | ::|:||
Db         905 TKGLGKTVP---LESESFVKWRDDVVVPCGKPVP-SSIRSSELMYNEYIVYNTSQVKMQFL 960

Qy         536 LEV 538
                 |:|
Db         961 LKV 963
```
```
RESULT 14
ID      Q9ZSV1          PRELIMINARY;     PRT;     980 AA.
AC      Q9ZSV1;
DT      01-MAY-1999 (TrEMBLrel. 10, Created)
DT      01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      POLY(ADP)-RIBOSE POLYMERASE (EC 2.4.2.30).
DE      PARP1.
OS      Zea mays (Maize).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae: PACC clade;
OC      Panicoideae; Andropogoneae; Zea.
OX      NCBI_TaxID=4577;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99026291; PubMed=9808734;
RA      Mahajan P.B., Zuo Z.;
RT      "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL      Plant Physiol. 118:895-905(1998).
DR      EMBL; AF093627; AAC79704.1; -.
DR      HSSP; P26446; 1A26.
DR      InterPro; IPR001357; BRCT.
DR      InterPro; IPR001290; PARP.
DR      InterPro; IPR004102; PARP_reg.
DR      InterPro; IPR003034; SAP.
DR      InterPro; IPR001510; Znf-PARP.
DR      Pfam; PF00533; BRCT; 1.
DR      Pfam; PF00644; PARP; 1.
DR      Pfam; PF02877; PARP_reg; 1.
DR      Pfam; PF00645; zf-PARP; 2.
DR      ProDom; PD004675; Znf-PARP; 2.
DR      SMART; SM00292; BRCT; 1.
DR      PROSITE; PS50172; BRCT; 1.
DR      PROSITE; PS50064; PARP_zN_FINGER_2; 2.
KW      Transferase; Glycosyltransferase.
SQ      SEQUENCE  980 AA;  110475 MW;  9D8AED26BC37E5C1 CRC64;

Query Match             20.3%;  Score 578;  DB 10;  Pred. No. 1.1e-35;  Length 980;
Best Local Similarity   30.6%;
Matches  166;  Conservative  96;  Mismatches  195;  Indels  86;  Gaps  21;

Qy          23 KKKGRQAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNI 82
                 | ||| || :|  | ||     |:  || ::::
Db         491 KVKGRSAVHESSGLQDTAHILE---------------------DGKSI---YNATLNMSDL 527

Qy          83 ENNNNKFYIIQLL-QDSNRFFTCWNRWGRVG--EVGQSKINHFTRLEDAKKDFEKKFREK 139
                 :  |||| :|| :|:||| | :  ||||| :  |||||||  | || :|  | ||
Db         528 ALGVNSYYVLQIIEQDDGSECYVFRKWGRVGSEKIGGQKLEEMSKTE-AIKEFKRLFLEK 586

Qy         140 TKNNWAE---RDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPAT 196
                 | :| |      ||| :|:|| :| |   :  ||    |:    |  |:|   |
Db         587 TGNSWEAWECKTNFRKQPGRFYPLDVD-------YGVKKAPKRKDISEMK-SSLAPQL 636

Qy         197 QKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQOIARGFEALEALEEALKGPTDGGQS 256
                 || |   |: | :   || :|| |||    | |  |:  ||| | | || || |
Db         637 LELMKMLFNVETYRAAMMEFEINMSEMPLGKLSKENIEKGFEALTEIQNLLKDTADQALA 696

Qy         257 LEE-----LSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAV-SE 310
                 |:      | ||| | |||  | || ||: || :| ||    :| ||:  |   :|
Db         697 VRESLIVAASNRFFTLIP----SIHPHIIRDEDDLMIKAKMLEALQDIEIASKIVGFDSD 752
```

```
Qy    311  QEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPT-------LQ 363
                 :::    ||  |  |  || |  :  : :||:||: || || | ||
Db    753  SDES------LDDRYMKLHCDITPLAHDSEDYKLIEQYL---LNTHAPTHKDWSLELE 801

Qy    364  HIWKVNQEGEEDRFQAH-SKIGNRKLLWHGTNMAVVAAILTSGLRIMPH----SGGRVGK 418
                 ||:: || ::: :: | :|: |||||  |  |    ||:||||   |
Db    802  EVFSLDRDGELNKYSRYKNNLHNKMLLWHGSRLTNFVGILSQGLRIAPPEAPVTGYMFGK 861

Qy    419  GIYFASENSKSAGYVIGMKCGA---HHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDS 475
                |:|||  |||||   : |||    ||  ||||  |   :    || ||||
Db    862  GLYFADLVSKSAQY-----CVVDRNNPVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS 915

Qy    476  VIARGHTEPDPTQDTELELDGQQVVPQGGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYL 535
                 |||  | || |  :  :::|    |||| ||  | ::||:|||||:|:: ::::|
Db    916  TKGLGKTVP---LESEFVKWRDDVVVPCGKPVP-SSIRSSELMYNEYIVYNTSQVKMQFL 971

Qy    536  LEV 538
                |:|
Db    972  LKV 974


RESULT 15
Q9XUA5
ID   Q9XUA5      PRELIMINARY;   PRT;   727 AA.
AC   Q9XUA5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC8.1 PROTEIN.
GN   AC8.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McMurray A.A.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1999).
DR   EMBL; Z83097; CAB05448.1; -.
DR   HSSP; P26446; 1A26.
DR   InterPro; IPR001290; PARP.
DR   InterPro; IPR004102; PARP_reg.
DR   InterPro; IPR001510; Znf-PARP.
DR   Pfam; PF00644; PARP; 1.
DR   Pfam; PF02877; PARP_reg; 1.
DR   Pfam; PF00645; zf-PARP; 1.
DR   ProDom; PD004675; Znf-PARP; 1.
DR   PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
SQ   SEQUENCE 727 AA; 82884 MW; 530ABA8E991FFEFD CRC64;


Query Match        18.7%;  Score 534;  DB 5;  Length 727;
Best Local Similarity  28.6%;  Pred. No. 1.8e-32;
Matches  162;  Conservative 104;  Mismatches 211;  Indels  90;  Gaps  20;

Qy     37  RSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQ 96
                ::|::| ::  |: :|    |||  | :   |: |:| ||  :||  |::
Db    182  KATGEYVEALAAKGG--STEPATPASASP-TPPEAETPVLSAEGSPESSNKRPASAEIIE 238

Qy     97  ------DSNRFF--TCWNRWGRVGEVGQSKI------------NHFTRLED 127
                      |: :|   |||  || ::| ||:         : :|
Db    239  IDGEGNPDENDFAKKRRMKKEARLMEVQKKRMKKQSDLLWEYRQIFERMPYTDNISILRE 298

Qy    128  AKKDF------EKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGP 179
                |  |      |||  | | || :|  |  :| || |:|:: || :  : |
Db    299  NEQDIPEGHDRTAQDVFHEKTKNDWIYRKHFRKMPGMFSVVETDYSE-----FVGTNNGH 353
```

```
Qy    180  VRTVT--KRVQPCS---LDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA 234
                 ::|   |  ||  :  ||: | |  ||  : | ||  || ||||  ||| | || |
Db    354  KKKITPGSKITPGSKTLLPKSVKEVVMSIFDVENMKSALKSFEIDVNKMPLGRLSHNQIN 413

Qy    235  RGFEALEAEEAL-KGPTDGGQSLEELSSHFYTVIPHNFGHSOPPPINSPELLQAKKDML 293
                || | |::|   ||| ||   |    ::|| | :||||||:   |  ||: |:| | ::
Db    414  LAFEVLNDISDLLVKLPIDASKIL-DFSNKFYTIIPHNFGMRVPEPIDSFHKIKEKNNML 472

Qy    294  LVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT 353
                |::| :|  |  : :: |  || :|: || |||| || || :|:| || |:|:: :
Db    473  NALLDIKFAYDQISGGDVPASTSLSIDPVDINYRKLKCIMEPLQQGCDDWNMIHQYLKNT 532

Qy    354  -GSNHRCPT-LQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH 411
                 ::|| |    |::||:|::::| :  ||: :|:|||||||| |||  |:|||||||||
Db    533  HGATHDLKVELIDILKVNRDNESSKFKRH--IGNRRLLWHGSGKMNFAGILGQGLRIAPP 590

Qy    412  ----SGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGR-EHHINTDNPSL 466
                     |||| |||||||: || |||| |:|    ||  || :  :||
Db    591  EAPVSGYMFGKGVYFADMFSKSFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNVSR 647

Qy    467  KSPPPGFDSVIARG--------HTEPD---PTQDTELELDGQQVVVPQGGQPVPCPEFS 513
                :: || || :|    ::  :   ::::| :
Db    648  QTLPAGFQSVQGVGRQCPREIGSYKNPDGYTVPLGLITYMQLQGKQNV----------- 694

Qy    514  SSTFSQSEYLIYQESQCRLRYLLEVHL 540
                || || : |::|||| :|||||: :|
Db    695  DYHLLYNEFIVYDVDQIQLKYLVRVKM 721
```

Search completed: August 29, 2002, 08:01:34
Job time: 366 sec

us-09-701-586b-6.rspt

Mon Sep 9 08:38:41 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:58:13 ; Search time 46.79 Seconds
(without alignments)
281.894 Million cell updates/sec

Title: US-09-701-586B-6
Perfect score: 2854
Sequence: 1 MSLLFLAMAPKPKPWVQTEG.........EYLIYQESQCRLRYLLEVHL 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	672	23.5	1014	4	US-09-078-347A-3	Sequence 3, Appli
2	670	23.5	1013	4	US-09-860-886-2	Sequence 2, Appli
3	138.5	4.9	1327	4	US-09-196-387-2	Sequence 2, Appli
4	113.5	4.0	1164	4	US-08-923-992A-2	Sequence 2, Appli
5	112.5	3.9	717	4	US-09-078-925-1	Sequence 1, Appli
6	112.5	3.9	1098	4	US-08-923-992A-8	Sequence 8, Appli
7	110.5	3.9	1128	4	US-08-923-992A-6	Sequence 6, Appli
8	109	3.8	743	2	US-08-910-925-3	Sequence 3, Appli
9	108.5	3.8	1871	3	US-08-694-869-1	Sequence 1, Appli
10	108.5	3.8	1871	4	US-09-349-546-1	Sequence 1, Appli
11	108.5	3.8	2548	4	US-09-172-422-1	Sequence 1, Appli
12	107.5	3.8	1164	4	US-08-923-992A-10	Sequence 10, Appli
13	105.5	3.7	1464	4	US-09-045-360-2	Sequence 2, Appli
14	105	3.7	641	1	US-08-961-083-160	Sequence 160, App
15	103.5	3.6	1528	1	US-08-326-117B-2	Sequence 2, Appli
16	103.5	3.6	1528	3	US-08-982-129-2	Sequence 2, Appli
17	102.5	3.6	2482	2	US-08-328-254-6	Sequence 6, Appli
18	101.5	3.6	1104	4	US-08-923-992A-4	Sequence 4, Appli
19	101.5	3.6	1461	2	US-08-993-228-10	Sequence 10, Appli
20	101.5	3.6	3075	2	US-08-460-309-5	Sequence 5, Appli
21	101.5	3.6	3075	2	US-08-125-077-5	Sequence 5, Appli
22	101	3.5	452	2	US-08-686-599A-18	Sequence 18, Appli
23	101	3.5	493	2	US-08-686-599A-5	Sequence 5, Appli
24	101	3.5	493	2	US-08-686-599A-16	Sequence 16, Appli
25	100	3.5	443	2	US-08-795-475-6	Sequence 6, Appli
26	100	3.5	467	2	US-08-686-599A-17	Sequence 17, Appli
27	100	3.5	737	1	US-08-185-432-2	Sequence 2, Appli
28	100	3.5	737	1	US-08-185-432-4	Sequence 4, Appli
29	100	3.5	1786	4	US-08-973-462-8	Sequence 8, Appli
30	99.5	3.5	3248	1	US-08-353-700-1	Sequence 1, Appli
31	99.5	3.5	3248	5	PCT-US95-16216-1	Sequence 1, Appli
32	98	3.4	1939	4	US-09-310-187A-1	Sequence 1, Appli
33	97.5	3.4	163	2	US-08-044-618-6	Sequence 6, Appli
34	96	3.4	2391	4	US-08-446-855A-2	Sequence 2, Appli
35	96	3.4	2391	1	US-09-150-741-2	Sequence 2, Appli
36	95	3.3	472	1	US-08-021-601-8	Sequence 8, Appli
37	95	3.3	472	8	US-08-082-849B-8	Sequence 8, Appli
38	95	3.3	472	5	PCT-US94-01624-8	Sequence 8, Appli
39	94.5	3.3	456	1	US-08-021-601-6	Sequence 6, Appli
40	94.5	3.3	456	8	US-08-082-849B-6	Sequence 6, Appli
41	94.5	3.3	456	5	PCT-US94-01624-6	Sequence 6, Appli
42	94	3.3	557	4	US-08-979-608A-5	Sequence 5, Appli
43	94	3.3	1066	4	US-09-541-782-8	Sequence 8, Appli
44	93	3.3	504	3	US-09-231-529-6	Sequence 6, Appli
45	93	3.3	504	4	US-08-977-816-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-078-347A-3
; Sequence 3, Application US/09078347A
; Patent No. 6132968
; GENERAL INFORMATION:
; APPLICANT: Le, Xiao-Chun
; APPLICANT: Weinfeld, Michael
; APPLICANT: Xing, James Z.
; TITLE OF INVENTION: Methods for Quantitating Low Level
; TITLE OF INVENTION: Modifications of Nucleotide Sequences
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,347A
; FILING DATE: 13-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UALB-03283
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1014 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-078-347A-3

Query Match 23.5%; Score 672; DB 4; Length 1014;
Best Local Similarity 33.6%; Pred. No. 1.1e-59;
Matches 187; Conservative 100; Mismatches 192; Indels 78; Gaps 23;

Qy 7 AMAPKPKPWVQTEGPEKKKGR-----QAGREEDP---FRSTAEALKAIPAEKRIIRVDPT 58

```
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  650-493-4935
;     TELEFAX:  650-493-5556
;     TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1013 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FRAGMENT TYPE:  internal
US-08-860-886-2

Query Match             23.5%;  Score 670;  DB 4;  Length 1013;
Best Local Similarity   33.6%;  Pred. No. 1.7e-59;
Matches  186;  Conservative  101;  Mismatches  195;  Indels  72;  Gaps  23;

Qy    7 AMAPKPKPWVQTEGPEKKKGR-----QAGREEDP---FRSTAEALKAIPAEKRIIRVDPT  58
Db  501 ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLE-------------  546

Qy   59 CPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWGRVGEV-G  115
Db  547 ------KGGKV--FSATLGLVDIVKGTNSYYKLQLLEDDKENRYW-IFRSWGRVGTVIG  596

Qy  116 QSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVK  174
Db  597 SNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVKK  653

Qy  175 VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA  234
Db  654 LTVNP-GTKSKLPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQ  706

Qy  235 RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL  294
Db  707 AAYSILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLD  765

Qy  295 VLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-  353
Db  766 NLLDIEVAYSLLRGGSDDSSKD----PIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTH  821

Qy  354 GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH  411
Db  822 ATTHNAYDLEVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPP  881

Qy  412 ---SGGRVGKGIYFASENSKSAGYVIGMKCGAHH-----VGYMFLGEVALGREHHINTDNP  464
Db  882 EAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILLGEVALGNMYELKHAS-  935

Qy  465 SLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYLI  524
Db  936 HISKLPKGKHSVKGLGKTTPDPS--ANISLDG-VDVPLGTGI-SSGVNDTSLLYNEYIV  990

Qy  525 YQESQCRLRYLLEV  538
Db  991 YDIAQVNLKYLLKL  1004


RESULT  3
US-09-196-387-2
;  Sequence 2, Application US/09196387
;  Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
```

```
Db  502 ALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLEHSAHVLE-------------  547

Qy   59 CPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQD--SNRFFTCWNRWGRVGEV-G  115
Db  548 ------KGGKV---FSATLGLVDIVKGTNSYYKLQLLEDDKENRYW-IFRSWGRVGTVIG  597

Qy  116 QSKINHFTRLEDAKKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAVVK  174
Db  598 SNKLEQMPSKEDAIEQFMKLYEEKTGNAWHSK-NFTKYPKKFYPLEIDYGQDE--EAVKK  654

Qy  175 VDRGPVRTVTKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIA  234
Db  655 LTVNP-GTKSKLPKP------VQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQ  707

Qy  235 RGFEALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLL  294
Db  708 AAYSILSEVQQAVSQGSSDSQIL-DLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLD  766

Qy  295 VLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQT-  353
Db  767 NLLDIEVAYSLLRGGSDDSSKD---PIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTH  822

Qy  354 GSNHRCPTLQ--HIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPH  411
Db  823 ATTHSAYDLEVIDIFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGILSQGLRIAPP  882

Qy  412 ---SGGRVGKGIYFASENSKSAGYVIGMKCGAHH-----VGYMFLGEVALGREHHINT  461
Db  883 EAPVTGYMFGKGIYFADMVSKSANY------YHTSQGDPIGLILLGEVALGNMYELKH  934

Qy  462 DNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSE  521
Db  935 AS-HISRLPKGKHSVKGLGKTTPDPS--ANISLDG--VDVPLGTGISSGVIDTSLL-YNE  988

Qy  522 YLIYQESQCRLRYLLEV  538
Db  989 YIVYDIAQVNLKYLLKL  1005


RESULT  2
US-08-860-886-2
;  Sequence 2, Application US/08860886
;  Patent No. 6335009
;  GENERAL INFORMATION:
;    APPLICANT:  Burkle, Alexander
;    APPLICANT:  zur Hausen, Harald
;    APPLICANT:  Jan-Heiner, Kupper
;    TITLE OF INVENTION:  VECTORS AND VIRUSES FOR USE
;    TITLE OF INVENTION:  IN GENE THERAPY
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/860,886
;      FILING DATE:  03-OCT-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  8484-0028-999
```

;   CITY: Hackensack
;   STATE: New Jersey
;   COUNTRY: USA
;   ZIP: 07601
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/196,387
;   FILING DATE:
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/095,225
;   FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Jackson Esq., David A.
;   REGISTRATION NUMBER: 26,742
;   REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 201-487-5800
;   TELEFAX: 201-343-1684
;   TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1327 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
US-09-196-387-2

```
Query Match        4.9%;    Score 138.5;  DB 4;   Length 1327;
Best Local Similarity  24.4%;  Pred. No. 0.00012;
Matches   73;  Conservative   41;  Mismatches 102;  Indels   83;  Gaps   16;

Qy   284 ELLQAKKDMLLVLADI---ELAQ-ALQAVSEQEKTVEEVPHPL-----DRDYQLLKCQLQ 334
             :: : :| | |||: ||:  | | ||:| : :| | |       | |||| ||
Db  1045 DIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQ 1104

Qy   335 ---LLDSGAP--EYKVIQTYLEQTGSNHRCP--------TLQHIWKVNQEGEEDRF- 377
            ||:|:  :| |: |:| :|  |  |||||       : || ||:| :|
Db  1105 GTILLDL-APEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFC 1163

Qy   378 ------QAHSKLGNRKLLWHGTNMAVVAAILTSGL-RIMPHSGGRVGKGIYFASENSKS 429
               :||:|:: | ||::| :| |||| ||:||  |  ||||||||||::||
Db  1164 HRQKEVSEENHNHNERMLFHGS--PFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKS 1221

Qy   430 AGYVIGM----KCGAH-------HVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVI 477
            ||| :     || || :  :    : ||: |: || : || :|:
Db  1222 NQYVYGIGGGTGCPTHKDRSCYICHRQMLFC-RVTLGKS-FLQFSTMKMAHAPPGHHSVI 1279

Qy   478 ARGHTEPDPTQDTELELDGQQVVVPGGQPVPCPEFSSSTFSQSEYLIYQESQCRLRYLL 536
              |:    |:|:|||||||||||:  :    :|||:: :|: :|||||
Db  1280 -------------------GRP------SVNGLAYAEYVIYRGEQAYPEYLI 1306
```

RESULT 4
US-08-923-992A-2
; Sequence 2, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
;   APPLICANT: Tai, Joseph Y.
;   APPLICANT: Blake, Milan S.
;   TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
;   TITLE OF INVENTION: Streptococcal Beta Antigens
;   NUMBER OF SEQUENCES: 34
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;   STREET: 1100 New York Avenue, N.W., Suite 600

;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20005
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/923,992A
;   FILING DATE: 05-SEP-1997
;   CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/024,707
;   FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Esmond, Robert W.
;   REGISTRATION NUMBER: 32,893
;   REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 371-2600
;   TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1164 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-923-992A-2

```
Query Match        4.0%;    Score 113.5;  DB 4;   Length 1164;
Best Local Similarity  20.2%;  Pred. No. 0.034;
Matches  107;  Conservative   72;  Mismatches 205;  Indels 147;  Gaps   25;

Qy    28 QAGREEDPPRSTABEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN 86
          || : :| :|  |||        :||:  :|  |:|  |  |       | :|
Db   309 QATQVKNQFLENAQKLKEIQ----------------PLIKETNVKLYKAMSESLEQVEKELKHN 356

Qy    87 NKFYIIQLLQDSNRFFTCW-----NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREK--- 139
          ||:  :|:  :  || ||     ||  |  |  :| | :  |::|  | :||:
Db   357 SEANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFKTSEQV 416

Qy   140 -------TKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPV---RTVT 184
                 ||| : |  :|:|  ||  |:|   | | | | | |  |:|    |:|:
Db   417 TPKKRVKRDLAAENENQQKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDFSDLL 473

Qy   185 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE 244
          |||    :|  ||:||:|    | |:|   |  |:      : : :  |:
Db   474 TKYNPSVSDRISTNYKTN-----TDNHKIAEITIKNL--KLNESQTV--------- 513

Qy   245 EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQA 304
          ||| |  | |:    :  |||   |: |   |||   |   |||:| | | ||::| |
Db   514 -TLKAKDDSGNVVEKT----FTITVQKKEEKQVP--KTPE----QKD--------- 549

Qy   305 LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQ-LLDSGAPEYKVIQTYLEQTGSNHRCPTL 362
          ||                |:| : | | |     |: | || | : | :| | |  :
Db   550 ------SKTEEKVPQEPKSND---KNQLQELIKSAQQELEKLEKAIKEL--------- 589

Qy   363 QHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYF 422
          | :  |:| :|  |:     |:|| ||:||||  ||| :||  |:  ||| |||| |:
Db   590 ------MEQPEIPSNPEYGIQKSIWESQKEPIQEAI-TSFKKIIGDSSSKYYTEHYF 639

Qy   423 ASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHIN--TDNPSLKSPPPGFDSVIAR- 479
          |:|:|  ||   |  |     :|  || :  | | :|    : |::| |  |   |:
Db   640 NKYKSDFMNYQL------HAQMEMLTRKVV---QYMNKYPDNAEIKKI---FESDMKRT 686

Qy   480 -----GHTEPD----------PTQDTELELDGQQVVVPGGQPVPCPEFS 513
                 :| ||          ||  |||| ||| ||  |  |  ||:
Db   687 KEDNYGSLENDALKGYFEKYFLTPFNKIKQIVDDLDKKVEQDQPAPIPENS 737
```

RESULT 5
US-08-910-925-1
; Sequence 1, Application US/08910925
; Patent No. 6162601
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,925
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0365 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 717 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: FIBRNOT01
; CLONE: 53219
US-08-910-925-1

Query Match   3.9%; Score 112.5; DB 4; Length 717;
Best Local Similarity 19.3%; Pred. No. 0.019;
Matches 102; Conservative 81; Mismatches 229; Indels 117; Gaps 21;

```
Qy   20 GPEKKGRQAGREEDP----FRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNC 75
            | |: :|:|| ||   |     | :|: |
Db   87 GGERTRRESRQESDPEDDDVKKPALQSSVVATSKERTRD-----------LIQDQN- 133

Qy   76 TLNQTNIENNNKFYII----QLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKK 130
            :|| ::| :| || :|   |  | ||   |  : :|:|  |  : |:
Db  134 -MDEKGKQRNRRIFGLLMGTLQKFKQEST---VATERQKRRQEIEQ-KLE--VQAEEERK 186

Qy  131 DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAE-DEAQEAVVKVDRGPVRTVTKR--- 186
            | :|: |: :| :|| | | :| |   :| :|   |:|   :|
Db  187 QVENERRELFEERRAKQTELRLLEQKVELAQLQEWNEHNAKIIKY-----IRTKTKPHLF 242

Qy  187 VQPCSLDPATQKLIINIFSK--EMPKNTMALMDLDVKKMPLG-----KLSKQQIARGFE 238
            ||||||  : | |  || :  |:  |     ||     |  |   |
Db  243 YIPGRMCPATQKLIEESQRKMNALFEGRRIEFAEQINKMEARPRRQSMKEKEHQVVRNEE 302

Qy  239 ALEALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLAD 298
            :|  |  |  |  | :| : | |  :| :|     : ||
Db  303 QKAEQEEGKVAQRE--EELEETGNQHNDVIEEAGEEE-----------EKEIAIVHSD 348

Qy  299 IELAQAIQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHR 358
            | ||  :|| ::| |:     | || :| | ::| :|
Db  349 AEKEQ-----EEEQKQEMEVKMEETEVRESEKQ-QDSQPEEVMDVLEMVE-------- 394

Qy  359 CPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGK 418
            :|:   :|     | |  |     | || || :|
Db  395 -NVKHVIADQEVMETNRVESVEPSENE------------ASKELEPE-------- 428

Qy  419 GIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIA 478
            :|: ||  |   :|: || |      | ::|:
Db  429 -MEFEIEPDK-------ECKSLSPGKENVSALDMEKESEEKEKESEPQPEP-------- 472

Qy  479 RGHTEPDPTQDTELELDGQQVVPQGQPVPCPEFSSSTFSQSEYLIYQE 527
            :|:    |:||    ||    || | :|   |  : :|:|| ||
Db  473 --VAQPQSQPQLQLQSQSQPVLQSQPPSQPEDLSLAVLQPTPQVTQE 519
```

RESULT 6
US-08-923-992A-8
; Sequence 8, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

Query Match   3.9%; Score 112.5; DB 4; Length 1098;
Best Local Similarity 20.0%; Pred. No. 0.039;
Matches 107; Conservative 75; Mismatches 199; Indels 153; Gaps 26;

```
Qy   28 QAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN 86
            : :: | : |||     :|  :
Db  272 QATQVKNQFLENAQKLKEIQ-----------PLIKETNVKLYKAMSESLEQVEKELKHN 319

Qy   87 NKFYIIQLLQDSNRFFTCW---NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREK--- 139
            : |:|| :| | || :|   |: ||:  |||:|:  :|
Db  320 SEANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFKTSEQV 379
```

```
Query Match                    3.9%;   Score 110.5; DB 4;   Length 1128;
Best Local Similarity         20.0%;   Pred. No. 0.066;
Matches 106; Conservative 73; Mismatches 205; Indels 147; Gaps 25;

Qy    28  QAGREEDFPRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN  86
          ||  :.||  :  |.|| :  :.||                  ::..|| :.| :|
Db   273  QATQVKNQFLENAQKLKEMQ-----------------PLIKETNVKLYKAMSESLEQVEKELKHN 320

Qy    87  NKFYIIQLLQDSNRFFTCW----NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREK---- 139
          :: |:||||:|  :   |   :| : |  |.:: ||||:   || :|||:
Db   321  SEANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFPKTSEQV 380

Qy   140  -----------TKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPV---RTVT  184
                      ||  |   || |   ||: ||| ||||:  : ||: :|   |:||
Db   381  TPKKRVKRDLAANENNQQKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDFSDLL  437

Qy   185  KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE  244
          :  | :     | |:|:|:     :::|  |:    :::||: ::       |||||||
Db   438  TKYNPSVSDRISTNYKTN------TDNHKIAEITIKNL--KLNESQTV-----------  477

Qy   245  EALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQA  304
          : | | : |.   ::| | |||:|||  :|| :|::  :|| :   |   ::||:|||
Db   478  -TLKAKDDSGNVVEKT-----FTITVQKKEEKQVP----KTPE-----QKD--------  513

Qy   305  LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQ-LLDSGAPEYKVIQTYLEQTGSNHRCPTL  362
          :     |:| :|| | |||| :|||||| | :|        || |: :|||:||||||||
Db   514  ------SKTEEKVPQEPKSND----KNQLQELIKSAQQELEKLEKAIKEL--------  553

Qy   363  QHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYF  422
          :|:|: | : |:  :: : || |::| :  |: || || | : |   |  | |:|||||
Db   554  -----------MEQPEIPSNPEYGIQKSIWESQKEPIQEAI-TSFKKIIGDSSSKYYTEHYF  603

Qy   423  ASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHIN--TDNPSLKSPPPGFDSVIAR-  479
          :  :      |::  : |  |||:|::::      :||  |  |  |::::|||
Db   604  NKYKSDFMNYQL------HAQMEMLTRKVV-----QYMNKYPDNAEIKKI---FESDMKRT 650

Qy   480  AR---------GHTEPD-----------PTQDTELELDGQQVVVPQGGQPVPCPEFS 513
           |         |||||            |||||||||||||||||||||||||||||
Db   651  KEDNYGSLENDALKGYFEKYFLTPFNKIKQIVDDLDKKVEQDQPAPIPENS 701
```

RESULT 8
US-08-910-925-3
; Sequence 3, Application US/08910925
; Patent No. 6162601
; GENERAL INFORMATION:
;   APPLICANT: Bandman, Olga
;   APPLICANT: Lal, Preeti
;   APPLICANT: Shah, Purvi
;   TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Incyte Pharmaceuticals, Inc.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/910,925
;   FILING DATE: Herewith
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:

RESULT 7
US-08-923-992A-6
; Sequence 6, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
;   APPLICANT: Tai, Joseph Y.
;   APPLICANT: Blake, Milan S.
;   TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
;   TITLE OF INVENTION: Streptococcal Beta Antigens
;   NUMBER OF SEQUENCES: 34
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;   STREET: 1100 New York Avenue, N.W., Suite 600
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20005
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/923,992A
;   FILING DATE: 05-SEP-1997
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/024,707
;   FILING DATE: 06-SEP-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Esmond, Robert W.
;   REGISTRATION NUMBER: 32,893
;   REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 371-2600
;   TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1128 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-923-992A-6

```
;   NAME: Billings, Lucy J.
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0365 US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  650-855-0555
;   TELEFAX:  650-845-4166
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 743 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GenBank
;   CLONE: 1684847
US-08-910-925-3
```

```
Query Match            3.8%;    Score 109; DB 4;  Length 743;
Best Local Similarity  20.0%;   Pred. No. 0.046;
Matches 115;  Conservative  79;  Mismatches 190; Indels 190;  Gaps    29;

Qy   20  GPEKKGRQAGREEDP---FRSTAEAALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNC 75
             |  :||: |::: ||       |  :|    ||  | ||        |
Db   86  GGERRTRRESRQESDPEDDVKPALQSSVVATSKERTRRD-----------LIQDQN- 132

Qy   76  TLNQTNIENNNKFYII----QLLQDSNRFFTCWNRWGRVGEVGQSKINHFTRLEDAKK 130
              :::  ||| |:    |    |:| ::   | | | |::  |::  :| :|
Db   133 -MDEKGKQRNRRIFGLLMGTLQKFKQEST---VATERQNRRQEIEQ-KLE-VQAEEERK 185

Qy   131 DFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQAE-DEAQEAVVKVDRGPVRTVTKR--- 186
             | | |: :|| || | |  |  |  | | |:     |::|| | |  | | ||
Db   186 QVENERRELFEERRAKQTELRLLEQKVELAQLQEWNEHNAKIIKY---IRTKTKPHLF 241

Qy   187 VQPCSLDPATQKLITNIFSKEMFKNTMALMD------LDVKKMPLG-----KLSKQQI 233
             |   |   || :    |   |  | || |       | |   || |       | | | |
Db   242 YIPGRMCPATQKLI-----EESQRKMNALFDGRRIEFAEQINKMEARPRRQSMKEKEHQV 296

Qy   234 ARGFEALEALEEALKGPTDGGQSLEEL----SSHFYTVIPHNFGHSQPPPINSPELLQA- 288
                |                            | |  | ::  |
Db   297 VRNEEHKAEQEEG----KVAQREELVETGNQH-----------------NDVEIEEAG 334

Qy   289 --KKDMLLVLADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQLLDSGAPEYKV 345
            :|  ||  : ||    | :: ::  ||| ::
Db   335 EEEEKEIGIVHSDAEKEQ------EEEEQKQEMEVKMEEETEVRESEKQ-QDSQPEEVMD 387

Qy   346 IQTYLEQTGSNHRCPTLQHIWKVNQEGEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSG 405
              ::|||
Db   388 VLEMVE----NVKHVIADQEVMETNRVES-------------------------------- 412

Qy   406 LRIMPHSGGRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPS 465
             | | | |           ::|  | |||
Db   413 --VEP-------SENEASK-------------ELEPEMEFIEPDK-E 437

Qy   466 LKSPPPGFDSVIA----------RGHTEPDP----TQDTELELDGQQVVVPQGQP 506
            ::| : : :  |: |         |     |
Db   438 CKSLSPGKENVSALDMEKESDEKEKEKESEPQPEPVAQPQAQSQPQLQSQSEPQPQLQP 497

Qy   507 VPC-PEFSSSTFSQSEYLIYQESQCRLRYLLEVH 539
             ::   |   |: : |||
Db   498 EPAQPQL------QSQPQIQLQSQCHA--VLQSH 523

RESULT  9
US-08-694-869-1
;  Sequence 1, Application US/08694869
;  Patent No. 5994123
;  GENERAL INFORMATION:
;  APPLICANT: Olszewski, N.
;  APPLICANT: Tzafrir, I.
;  APPLICANT: Somers, D. A.
;  APPLICANT: Lockhart, B.
```

```
;   APPLICANT: Torbert, K.
;   TITLE OF INVENTION: SUGARCANE BACILLIFORM VIRUS
;   TITLE OF INVENTION: PROMOTER
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
;   STREET: P.O. Box 2938
;   CITY: Minneapolis
;   STATE: MN
;   COUNTRY: USA
;   ZIP: 55402
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/694,869
;   FILING DATE: 09-AUG-1996
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Woessner, Warren D
;   REGISTRATION NUMBER: 30,440
;   REFERENCE/DOCKET NUMBER: 600.369US1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 612-373-6900
;   TELEFAX: 612-339-3061
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1871 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-694-869-1
```

```
Query Match            3.8%;    Score 108.5; DB 2;  Length 1871;
Best Local Similarity  21.0%;   Pred. No. 0.25;
Matches 100;  Conservative  64;  Mismatches 135; Indels 177;  Gaps    25;

Qy   40   AEALKAIPAEKRIIRVDPTCPLSSNPGTQVVEDYNCTLNQTNIENNNKFYIIQLLQDSN 99
              |||:||: ::                 :|  |||: ::|| :::|
Db   919  AEAVK--PPEKK----------------SNYELLAKQLLIENSKLKMEKEILIEELN 957

Qy   100  RFFTCWNRWGR-----VGEVGQSKINHFTRLEDAKKDFEKKFREKTKNNWA---------- 145
             ::  : |||                  | :|::| |     :|
Db   958  KEIKAHQETKKGKELYIEEASTEVENEIETWKSRAELFEALYNEEVKKNKASTSSVTEGM 1017

Qy   146  --ERDHFVSHPGKYTLIEVQA-----------------EDEAQEAVV----- KVDRGPVRTVTKR 186
               |  |  || |  |                        ::|::||             ::||
Db   1018 YQVQIDHL-----RKELREVEATLEVNKVEESEEEAEEVMMASAVKDEMYRFPVIIEVPE 1072

Qy   187  VQPCSLDP-----ATQKLITNIFSKEMF----------------KNTMALMDLDVK----------- 221
             |                 :  ||:|                     |
Db   1073 VGKVQLTALLDTGATRSCINQVFIEEKFLQPTKFKVKIHGVNSVTKLDRQVKDGAKLWAG 1132

Qy   222  ---KMPL--GKL---SKQQIARGFEALEALEALEEALKGPTDGGQSLEELSSHFYTVIPHN 271
               :|:  || |         |  :| :::     |::|
Db   1133 ENWFRLPITYVGPMYMGEKTQMLIGCNFMQSLA---------GGVRLEGRTVTFYKYI---- 1181

Qy   272  FGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQE----------- KTVEEVPH 320
             |||                |||          :|:            :|::
Db   1182 ----------ASIKANEYLQAEAEEILV------------ATSEQEFINRSFMSKNKRLLEEMK- 1223

Qy   321  PLDRDY----------QLLKCQLQLLDSGAPEYKVIQTLEQTGSNHRCPTLQHIWKVN 369
             || :    |                    || ::::  | |  | :|
Db   1224 --EQGYMGEDTLAHWNKNQICKCKIELRN---PDLII------KDKPQTLLNIQK-- 1266
```

```
Qy  370 QEGEEDRFQAH-SKLGNRKLLW-----HGTNMAVVAAILTSGLRIMPHSGGRV-GK 418
             ::||:  |  ||    |  :||  ||   |  :||  ||  ||:  : ||
Db 1267 ---KEAMRKHIDALLERKVIRPSKSPHRTN----AFIVESGTSIDPKTGKEIRGK 1314
```

RESULT 10
US-09-349-546-1
; Sequence 1, Application US/09349546
; Patent NO. 6093569
; GENERAL INFORMATION:
;   APPLICANT: Olszewski, N.
;   APPLICANT: Tzafrir, I.
;   APPLICANT: Somers, D.A.
;   APPLICANT: Lockhart, B.
;   APPLICANT: Torbert, K.
;   TITLE OF INVENTION: Sugarcane bacilliform virus promoter
;   FILE REFERENCE: 600.369US2
;   CURRENT APPLICATION NUMBER: US/09/349,546
;   CURRENT FILING DATE: 1999-07-08
;   EARLIER APPLICATION NUMBER: US 08/694,869
;   EARLIER FILING DATE: 1996-08-09
;   EARLIER APPLICATION NUMBER: PCT/IB97/01338
;   EARLIER FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 1
;   LENGTH: 1871
;   TYPE: PRT
;   ORGANISM: sugarcane bacilliform virus
US-09-349-546-1

Query Match   3.8%; Score 108.5; DB 3; Length 1871;
Best Local Similarity   21.0%; Pred. No. 0.25;
Matches   100; Conservative 64; Mismatches 135; Indels 177; Gaps 25;

```
Qy   40 AEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIENNNKFYIIQLLQDSN 99
Db  919 AEAVK---PPEKK-------------------------SNYELLAKQLLIENSKLKMEKEILIEELN 957

Qy  100 RFFTCWNRWGR------VGEVGQSKINHFTRLEDAKKDFEKKFREKTKNNWA-------- 145
Db  958 KEIKAHQETKKGKELYIEEASTEVENEIETWKSRAELFEALYNEEVKKNKASTSSVTEGM 1017

Qy  146 ---ERDHFVSHPGKYTLIEVQA------EDEAQEAVV------KVDRGPVRTVTKR 186
Db 1018 YQVQIDHL-----RKELREVEATLEVNKVEESEEEAEEVMMASAVKDEMYRFPVIIEVPE 1072

Qy  187 VQPCSLDP-----ATQKLITNIFSKEMF------------KNTMALMDLDVK------- 221
Db 1073 VGKVQLTALLDTGATRSCINQVFIEEKFLQPTKFKVKIHGVNSVTKLDRQVKDGAKLWAG 1132

Qy  222 ---KMPL--GKL--SKQQIARGFEALEALREALKGPTDGGQSLEELSSHFYTVIPHN 271
Db 1133 ENWFRLPITYVGPMYMGEKTQMLIGCNFMQSLA--------GGVRLEGRTVTFYKYI--- 1181

Qy  272 FGHSQPPPINSPELLQAKKDMLLVLADIELAQALQAVSEQE--------KTVEEVPH 320
Db 1182 ------ASIKANEYLQAEAEEILV----------ATSEQEFINRSFMSKNKRLLEEMK- 1223

Qy  321 PLDRDY------QLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWKVN 369
Db 1224 --EQGYMGEDTLAHWNKNQIKCKIELRN---PDLII-------KDKPQTLLNIQK-- 1266

Qy  370 QEGEEDRFQAH-SKLGNRKLLW-----HGTNMAVVAAILTSGLRIMPHSGGRV-GK 418
Db 1267 ---KEAMRKHIDALLERKVIRPSKSPHRTN----AFIVESGTSIDPKTGKEIRGK 1314
```

RESULT 11
US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485

; GENERAL INFORMATION:
;   APPLICANT: Adams, Arwen E.
;   APPLICANT: Chiu, Choi Ying
;   APPLICANT: Duhl, David
;   APPLICANT: Gorman, Susan W.
;   APPLICANT: Leng, Song
;   APPLICANT: Sheffield, Val
;   APPLICANT: Welch, Juliet
;   TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
;   TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
;   TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
;   FILE REFERENCE: 200130.442
;   CURRENT APPLICATION NUMBER: US/09/172,422A
;   CURRENT FILING DATE: 1998-10-14
;   NUMBER OF SEQ ID NOS: 3
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 1
;   LENGTH: 2548
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-172-422-1

Query Match   3.8%; Score 108.5; DB 4; Length 2548;
Best Local Similarity   21.0%; Pred. No. 0.42;
Matches   88; Conservative 63; Mismatches 126; Indels 143; Gaps 25;

```
Qy   41 EALKAIPAEK-RIIRVDPTCPLSSNPGTQVYEDYNCTLNQTIENNNKFYIIQLLQDSN 99
Db  757 EILQRCKEEKYSITRKNPRTPLSDLQG--------MNALNEKNQHDTFDI--------- 798

Qy  100 RFFTCWNRWGRVGEVGQSKINHFTRLEDAKKDFEKKFREKTKNNWAERDH-------FV 151
Db  799 ----AWN--GRTG-IRQSRLSSGTSLLDK---DGIFANSTSSKLLERAHGILTRNKNFK 847

Qy  152 SHPG--KYTLIEVQA------------------EDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQ-- 197
Db  848 SKPALPKH-LLEVNSLKHLTRLTLQDRITKSLLH---KKKKPPSISAQFQASL 898

Qy  198 -KLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALE--EALKGPTDGG 254
Db  899 SKLMETLGQAEPY--FVKCIRSNAEKLPL-RFSDVLVLRQLRYTGMLETVQIRQSGYSSK 955

Qy  255 QSLEELSSHFYIVIPHNFGHSQPPPIN---------SPELLQAKKDMLLVLADIELAQA 304
Db  956 YSFQDFVSHFHVLLPRNI---IPSKFNIQDFFRKINLNPDNYQVGKTMVF-------- 1002

Qy  305 LQAVSEQEKT-VEEVPHP-----------LDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGS 355
Db 1003 --LKEQERQHLQDLLHQEVLRRIILLQRWFVLLCRQHFL--------------HLRQAS- 1046

Qy  356 NHRCPTLQHIWK--VNQEG-----EEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLR 407
Db 1047 ----VIQRFWRNYLNQKQVRDAAVQKDAFV-----------MASAAALLQASWR 1086
```

RESULT 12
US-08-923-992A-10
; Sequence 10, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
;   APPLICANT: Tai, Joseph Y.
;   APPLICANT: Blake, Milan S.
;   TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
;   TITLE OF INVENTION: Streptococcal Beta Antigens
;   NUMBER OF SEQUENCES: 34
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;   STREET: 1100 New York Avenue, N.W., Suite 600
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20005

```
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/923,992A
    FILING DATE: 05-SEP-1997
    CLASSIFICATION: 536
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 60/024,707
    FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
    NAME: Esmond, Robert W.
    REGISTRATION NUMBER: 32,893
    REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (202) 371-2600
    TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
    LENGTH: 1164 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-923-992A-10
```

```
Query Match           3.8%;   Score 107.5;  DB 4;  Length 1164;
Best Local Similarity 20.4%;  Pred. No. 0.14;
Matches 108;  Conservative 68;  Mismatches 210;  Indels 143;  Gaps 24;

Qy   28 QAGREEDPPRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQTNIE-NNN 86
         :: |   |  ||     |  ||  | |  || :|: |  | ::| | | :|||
Db  309 QATQVKNQFLENAQKLKEIQ----------PLIKETNVKLYKAMSESLEQVEKQLKHN 356

Qy   87 NKFYIIQLLQDSNRFFTCW----NRWGRVGEVGQSKINHFTRLEDAKKDFEKKFR---EK 139
         ::| ::|   | : ||       | | | ::| ||| ||  ||:|| |   | :
Db  357 SQANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKFKTSEQV 416

Qy  140 TKNNWAERDHFVSHPGKYTLIEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKL 199
         || ||  |  |   |  | | :|   |   |  |
Db  417 TPKKRVKRD-LAANENNQQKIELTVSPEN----ITVYEG------------EDL 453

Qy  200 ITNIFSKEMFKNTMALMDLDVKKMP--------LGKLSKQQIA---RGFEALEALEE 245
         | |  ||       |  | |  |           | :| ||       || |  |
Db  454 KFTLTAKSDSKTTLDFSDLLTKYNPSVSDRISTNYKTNTDNHKIAEITIKNLKLNESQTV 513

Qy  246 ALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQAL 305
         :|     |     |   ||:||      || :   |  |
Db  514 TLKAKDDSGNVVQKT----FTITVQKKEEKQVP--KTPE----QKD---------- 549

Qy  306 QAVSEQEKTVEEVP-HPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQH 364
         |   : |  | : |    |  |  |:
Db  550 ------SKTEEKVPQEPKSND----KNQLQEL------IKSAQQQLEKLEK---------- 584

Qy  365 IWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFAS 424
         |   :  :| || ::| | | | :|| |    | |:
Db  585 --AIKELMEQPEIPSNPEYGIQKSIWESQKEPIQEAI-TSFKKIIGDSSSKYYTEHYFNK 641

Qy  425 ENSKSAGYVIGMKCGAHHVGYMFLGEVALGREHHIN--TDNPSLKSPPPGFDSVIAR--- 479
          :   : |   |      |   || |     || |  | |  |:|    ||
Db  642 YKSDFMNYQL------HAQMEMLTRKVV---QYINKYPDNAEIKKI---FESDMKRTKE 688

Qy  480 ---GHTEPD------PTQDTELELDGGQVVVPGGQPVPCPEFS 513
             |  |        || :|  |   |  || |
Db  689 DNYGSLENDALKGYFEKYFLTPFNKIKQIVDDLDKKVEDQPAPIPENS 737
```

RESULT 13
US-09-045-360-2
; Sequence 2, Application US/09045360
; Patent No. 6207880

```
GENERAL INFORMATION:
    APPLICANT: Jens Kossmann
    APPLICANT: Ruth Lorberth
    TITLE OF INVENTION: PLANTS WHICH SYNTHESIZE A MODIFIED STARCH,
    TITLE OF INVENTION: PROCESS FOR THE PRODUCTION THEREOF AND MODIFIED STARCH
    NUMBER OF SEQUENCES: 4
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: FISH & NEAVE
        STREET: 1251 Avenue of the Americas
        CITY: New York
        STATE: New York
        COUNTRY: USA
        ZIP: 10020
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/09/045,360
        FILING DATE:
        CLASSIFICATION:
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: PCT/EP96/04109
        FILING DATE: 19-SEP-1996
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: DE 19547733.2
        FILING DATE: 20-DEC-1995
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: DE 19534759.5
        FILING DATE: 19-SEP-1995
    ATTORNEY/AGENT INFORMATION:
        NAME: Haley Jr., James F.
        REGISTRATION NUMBER: 27,794
        REFERENCE/DOCKET NUMBER: GFB-5
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: 212-596-9000
        TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
    LENGTH: 1464 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-360-2
```

```
Query Match           3.7%;   Score 105.5;  DB 4;  Length 1464;
Best Local Similarity 20.8%;  Pred. No. 0.34;
Matches 91;  Conservative 73;  Mismatches 193;  Indels 81;  Gaps 18;

Qy   44 KAIPAE-KRIIRVDPTCPLSSNPGTQVYEDYNCTLN--QTNIE---NNNNKFYIIQLLQ 96
         | |||  |:      :||   :: ||        |   ||:|      ||||||:
Db   61 KKIPMEKKRAFSSSPHAVLTDTSSELAEKFSLGGNIELQVDVRPPTSGDVSFVDFQVTN 120

Qy   97 DSNRFFTCWNRWGRVGEVGQSKINHFT-RLEDAKKDFEKKFPREKTKNNWAERDHFVSHPG 155
         | : | | |   | ::| | |   |   :   : | |   |
Db  121 GSDKLFLHW--------GAVKFGKETWSLPNDRPDGTKVYKNK----ALRTPFVK-SG 165

Qy  156 KYTLIEVQAEDEAQEAV------VKVDRGPVRTVTKRVQPCSLDPATQKLITNI 203
          ::| | : | | : ||      ||||||||:  :    |: |     |
Db  166 SNSILRLEIRDTAIEAIEFLIYDEAHDKWIKNNGGNFRVKLSRKEIRGPDVSVPEELVQI 225

Qy  204 FS-------KEMFKNTMALMDLDVKKMPLGKLSKQQIARGFEALEALEEALKGPTDGGQ 255
         :|       |:|| :   :     :|: |||||| | : : :||   :  :  |
Db  226 QSYLRWERGKQNYPPEKEKEYEARTVL---QEEIARG-ASIQDIRARLTKTNDKSQ 280

Qy  256 SLEELSSHFYTVIPHNFGHSQ----------PPPINSPELLQAKKDMLLVL-ADIEL 301
         | ||   | :|  | :| ||          ||| |||  |  |:|:|::|   |:|
Db  281 SKEEPLHVTKSDIPDDLAQAQAYIRWEKAGKPNYPPEKQIELEEARRELQLELEKGITL 340

Qy  302 AQALQAVSEQEKTVEVPHPLDRDYQLLKCQLQLLDSG----------APEYKVIQTYLEQT 353
         | || |  :|  || |      : | ::       ||           |  | :|||:| |
```

```
Db  341 DELRKTITKGEIKTKVEKHLKRSSFAVERIQRKKRDFGHLINKYTSSPAVQV-QKVLEEP 399
Qy  354 GSNHRCPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSG 413
Db  400 --------PALSKIKLYAKEKEE--QIDDPILNKKIFKVDDGELLVLVAKSSG-KTKVHLA 449
Qy  414 GRVGKGIYFASENSKSAG 431
Db  450 TDLNQPTLHWALSKSPG 467
```

RESULT 14
US-08-961-083-160
; Sequence 160, Application US/08961083
; Patent No. 6159469
GENERAL INFORMATION:
.; APPLICANT: Choi et. al.
.; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and vaccines
.; NUMBER OF SEQUENCES: 452
.; CORRESPONDENCE ADDRESS:
.;   ADDRESSEE: Human Genome Sciences, Inc.
.;   STREET: 9410 Key West Avenue
.;   CITY: Rockville
.;   STATE: Maryland
.;   COUNTRY: USA
.;   ZIP: 20850
.; COMPUTER READABLE FORM:
.;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
.;   COMPUTER: HP Vectra 486/33
.;   OPERATING SYSTEM: MSDOS version 6.2
.;   SOFTWARE: ASCII Text
.; CURRENT APPLICATION DATA:
.;   APPLICATION NUMBER: US/08/961,083
.;   FILING DATE:
.; PRIOR APPLICATION DATA:
.;   APPLICATION NUMBER:
.;   FILING DATE:
.; ATTORNEY/AGENT INFORMATION:
.;   NAME: Brookes, A. Anders
.;   REGISTRATION NUMBER: 36,373
.;   REFERENCE/DOCKET NUMBER: PB340P2
.; TELECOMMUNICATION INFORMATION:
.;   TELEPHONE: (301) 309-8504
.;   TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 160:
.; SEQUENCE CHARACTERISTICS:
.;   LENGTH: 641 amino acids
.;   TYPE: amino acid
.;   STRANDEDNESS: single
.;   TOPOLOGY: linear
.; MOLECULE TYPE: protein
US-08-961-083-160

```
Query Match         3.7%;  Score 105; DB 4; Length 641;
Best Local Similarity   19.0%;  Pred. No. 0.093;
Matches  96;  Conservative  67;  Mismatches 168;  Indels 174;  Gaps 21;

Qy   23 KKKGRQAGR--EEDPPRSTAEALKAIPAEKRIIRVDPTCPLSSNPGTQVYEDYNCTLNQT 80
           ::: :: ::    ::  ||      ::: |:    :: || ||
Db   43 KKKAEDAQKKYEDDQKRTEEKARKEAEASQKLNDVALVV------QNAYKEYREVQNQR 95
Qy   81 NIENNNNKFYIIQLLQDSNRFFTCWNRWGRVGEVGQSKINHF-TRLEDAKK---DFEKKF 136
           :: ::      ::|      |        :::||:    :: | :::||
Db   96 SKYKSDAEY--------------QKKLTEVDSKIEKARKEQQDLQNKF 129
Qy  137 REKTKNNWAERDHFVSHPGKYTLIEVQAED-EAQEAVVKVDRGPVRTVTKRVQPCSLDPA 195
           || ::    :: :||    ::|::   | |::   ||
Db  130 NE-----VRAVVVPEPNALAETKKKAEEAKAEEKVAK--------RKYDYA 167
Qy  196 TQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLS-----KQQIARGFEALEALEALKG 249
           |:       :: |:: :    |
```

```
Db  168 TLKV--------ALAKKEVEAKELEIEKLQYEISTLEQEVATAQHQVDNLKKLLAG 215
Qy  250 --PTDGGQSLEELSSHFYTVIPHNFGHSQPPINSPELLQAKKDMLLVLADIELAQALQA 307
Db  216 ADPDDGTEVIEAKLKK--------GEAE--LNAKQAELAKK------QTELEKLLDS 256
Qy  308 VSEQEKTVEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTGSNHRCPTLQHIWK 367
Db  257 LDPEGKTQDE----LDKEAEEAE----LDKKADELQNKVADLEKEISN------LEILL 301
Qy  368 VNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 427
Db  302 GGADXEDDTAALQNKLATKKAELEKTQKEDAALNELG----PD------341
Qy  428 KSAGYVIGMKCGAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPT 487
Db  342 --------GDEEETPAPAPQPEQPAPA--------PKPE 364
Qy  488 QDTELELDGQQVVPQ-GQPVPCPE 511
Db  365 QPAPAPKPEQPAPAPKPEQPAPAPK 389
```

RESULT 15
US-08-326-117B-2
; Sequence 2, Application US/08326117B
; Patent No. 5693491
GENERAL INFORMATION:
.; APPLICANT: JI, TAE
.; APPLICANT: BULLA, LEE A.
.; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
.; TITLE OF INVENTION: TOXIN
.; NUMBER OF SEQUENCES: 26
.; CORRESPONDENCE ADDRESS:
.;   ADDRESSEE: MORRISON & FOERSTER
.;   STREET: 2000 Pennsylvania Ave. N.W.
.;   CITY: Washington
.;   STATE: DC
.;   COUNTRY: USA
.;   ZIP: 20006-1812
.; COMPUTER READABLE FORM:
.;   MEDIUM TYPE: Floppy disk
.;   COMPUTER: IBM PC compatible
.;   OPERATING SYSTEM: PC-DOS/MS-DOS
.;   SOFTWARE: PatentIn Release #1.0, Version #1.30
.; CURRENT APPLICATION DATA:
.;   APPLICATION NUMBER: US/08/326,117B
.;   FILING DATE: 19-OCT-1994
.;   CLASSIFICATION: 435
.; ATTORNEY/AGENT INFORMATION:
.;   NAME: MILLMAN, ROBERT A
.;   REGISTRATION NUMBER: 36,217
.;   REFERENCE/DOCKET NUMBER: 7112-0037.00
.; TELECOMMUNICATION INFORMATION:
.;   TELEPHONE: (202) 887-1500
.;   TELEFAX: (202) 887-0763
.;   TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
.; SEQUENCE CHARACTERISTICS:
.;   LENGTH: 1528 amino acids
.;   TYPE: amino acid
.;   TOPOLOGY: linear
.; MOLECULE TYPE: protein
US-08-326-117B-2

```
Query Match         3.6%;  Score 103.5; DB 1; Length 1528;
Best Local Similarity   20.3%;  Pred. No. 0.58;
Matches 120;  Conservative  67;  Mismatches 207;  Indels 197;  Gaps 31;

Qy   65 PGTQVYEDYN----------CPTLNQTNIENNNNKFYIIQLLQDSNRFFTCWNRWGRVGE 113
           || :  :     |||        :: ||:|||||
Db  142 PGMQQYM-FNVRVDGQSLVAGVSLAIVNIDDNAP---IIQ------NFEPC------RVPE 186
```

```
Qy   114  VGQSKINHFT-----------------RLEDAKKDFEKKFREKTK--NNWAERD------ 148
               :|: .|           |.::   |.|:: | :|: |           |
Db   187  LGEPGLTECTYQVSDADGRISTEFMTFRIDSVRGDEETFYIERTNIPNQWMWLNMTIGVN 246

Qy   149  ---HFVSHPGKYTLLEVQAEDEAQEAVVKVDRGPVRTVTKRVQPCSLDPATQKLITNIFS 205
             :||:|:  | |       | :||   || || .::: : |:  .| :|:
Db   247  TSLNFVTSP--LHIFSVTALDSLPN-------THTVTMMVQVANVNSRPPRWL-EIFA 294

Qy   206  KEMFKN-------TMALMDLDVK-KMPLGKLSKQQIARGFEALEALEEALKG------PTD 252
             :|:      |: :|| .|:|:          | ::||            |    |
Db   295  VQQFEEKSYQNFTVRAIDGDTEINMPINYRLITNEEDTFFSIEALPGGKSGAVFLVSPID 354

Qy   253  GGQSLEEL-----------SSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVL 296
                    :|                 :| .|||    |.   :||: |         |:
Db   355  RDTLQREVFPLTIVAYKYDEEAFSTSTNVVIIVTDINDQRPEPIHKEYRLAIMEETPLTL 414

Qy   297  -------ADIELAQALQAVSEQEKTVEEVPHPLDRDYQLLKCQLQ-----------LL 336
                     |:  :| .| |           || || .|   :   |            :|
Db   415  NFDKEFGFHDKDLGQNAQYTVRLESV--DPPGAEAFYIAPEVGYQRQTFIMGTLNHSML 472

Qy   337  DSGAPEYKVIQTYLEQTGSN--HRCPTLQHIWKVNQEGEDRFQ----------- 378
             | |:::.| :|::|  .:|: |       | :| :|  |:  .:
Db   473  DYEVPEFQSITIRVVATDNNDTRHVGVALVHIDLINWNDEQPIFEHAVQTVTFDETEGEG 532

Qy   379  -----AHSK-LGN--RKLLWHGTNMAVVAAILTSGLRIMPHSGGRVGKGIYFASENS 427
                          |:  |:   :| ||  :|:  |:  :|:
Db   533  FFVAKAVAHDRDIGDVVEHTLLGNAVNFLTIDK-LTGDIRV-------------SAND 576

Qy   428  KSAGYVIGMKCGAHHVGYMFLGEVA---LGREHH------------INTDNPSLKSP--P 470
             |      |     .| .| ||   | .||:|         ||  |:||::|: .|  |
Db   577  -SFNY--HRESELFVQVRATDTLGEPFHTATSQLVIRLNDINNTPPTLRLPRGS 627

Qy   471  PGFDSVIARGH-------TEPDPTQDTELELD-GQQVVVPQGQPVPCPEF 512
             |.|: ||      |:||||:   |::             ||.
Db   628  PQVEENVPDGHVITQELRATDPDTTADLRFEINWDTSFATKQGRQANPDEF 678
```

Search completed: August 29, 2002, 07:58:18
Job time: 295 sec